



#5

SEQUENCE LISTING

<110> AGY Therapeutics  
Melcher, Thorsten  
Mueller, Sabine  
Chin, Daniel

<120> USE OF PROTEIN BIOMOLECULAR TARGETS IN THE TREATMENT AND VISUALIZATION  
OF BRAIN TUMORS

<130> 263/180 -- PEagleman -- AGY

<140> 09/983,000

<141> 2001-10-17

<160> 35

<170> PatentIn version 3.1

<210> 1  
<211> 3091  
<212> DNA  
<213> Homo sapiens

<220>  
<221> gene  
<222> (1)..(3091)  
<223> FTP-zeta SM1 exon 9 variant

<220>  
<221> VARSPLIC  
<222> (1262)..(1262)  
<223> Alternative splice site

<220>  
<221> misc feature  
<222> (1273)..(3091)  
<223> 3' Untranslated Region

<220>  
<221> CDS  
<222> (148)..(1272)  
<223>

Met Arg Ile Leu Lys Arg Phe Leu Ala  
1 5

tgc att cag ctc ctc tgt gtt tgc cgc ctg gat tgg gct aat gga tac 222  
Cys Ile Gln Leu Leu Cys Val Cys Arg Leu Asp Trp Ala Asn Gly Tyr  
10 15 20 25

tac aga caa cag aga aaa ctt gtt gaa gag att ggc tgg tcc tat aca 270  
Tyr Arg Gln Gln Arg Lys Leu Val Glu Glu Ile Gly Trp Ser Tyr Thr  
30 35 40

gga gca ctg aat caa aaa aat tgg gga aag aaa tat cca aca tgt aat 318  
Gly Ala Leu Asn Gln Lys Asn Trp Gly Lys Lys Tyr Pro Thr Cys Asn  
45 50 55

agc cca aaa caa tct cct atc aat att gat gaa gat ctt aca caa gta 366  
Ser Pro Lys Gln Ser Pro Ile Asn Ile Asp Glu Asp Leu Thr Gln Val  
60 65 70

aat gtg aat ctt aag aaa ctt aaa ttt cag ggt tgg gat aaa aca tca 414  
Asn Val Asn Leu Lys Lys Leu Lys Phe Gln Gly Trp Asp Lys Thr Ser  
75 80 85

ttg gaa aac aca ttc att cat aac act ggg aaa aca gtg gaa att aat 462  
Leu Glu Asn Thr Phe Ile His Asn Thr Gly Lys Thr Val Glu Ile Asn  
90 95 100 105

ctc act aat gac tac cgt gtc agc gga gga gtt tca gaa atg gtg ttt 510  
Leu Thr Asn Asp Tyr Arg Val Ser Gly Gly Val Ser Glu Met Val Phe  
110 115 120

aaa gca agc aag ata act ttt cac tgg gga aaa tgc aat atg tca tct 558  
Lys Ala Ser Lys Ile Thr Phe His Trp Gly Lys Cys Asn Met Ser Ser  
125 130 135

gat gga tca gag cat agt tta gaa gga caa aaa ttt cca ctt gag atg 606  
Asp Gly Ser Glu His Ser Leu Glu Gly Gln Lys Phe Pro Leu Glu Met  
140 145 150

caa atc tac tgc ttt gat ggc gac cga ttt tca agt ttt gag gaa gca 654  
Gln Ile Tyr Cys Phe Asp Ala Asp Arg Phe Ser Ser Phe Glu Glu Ala  
155 160 165

gtc aaa gga aaa ggg aag tta aga gct tta tcc att ttg ttt gag gtt 702  
Val Lys Gly Lys Gly Lys Leu Arg Ala Leu Ser Ile Leu Phe Glu Val  
170 175 180 185

ggg aca gaa gaa aat ttg gat ttc aaa ggc att att gat gga gtc gaa 750  
Gly Thr Glu Glu Asn Leu Asp Phe Lys Ala Ile Ile Asp Gly Val Glu  
190 195 200

Leu Asn Leu Leu Ile Asn Ser Thr Arg Lys Tyr Tyr Ile Tyr Asn Gly

220	225	230	
tca ttg aca tct cct ccc tgc aca gac aca gtt gac tgg att gtt ttt			894
Ser Leu Thr Ser Pro Pro Cys Thr Asp Thr Val Asp Trp Ile Val Phe			
235	240	245	
aaa gat aca gtt agc atc tct gaa agc cag ttg gct gtt ttt tgt gaa			942
Lys Asp Thr Val Ser Ile Ser Glu Ser Gln Leu Ala Val Phe Cys Glu			
250	255	260	265
gtt ctt aca atg caa caa tct ggt tat gtc atg ctg atg gac tac tta			990
Val Leu Thr Met Gln Gln Ser Gly Tyr Val Met Leu Met Asp Tyr Leu			
	270	275	280
caa aac aat ttt cga gag caa cag tac aag ttc tct aga cag gtg ttt			1038
Gln Asn Asn Phe Arg Glu Gln Gln Tyr Lys Phe Ser Arg Gln Val Phe			
	285	290	295
tcc tca tac act gga aag gaa gag att cat gaa gca gtt tgt agt tca			1086
Ser Ser Tyr Thr Gly Lys Glu Glu Ile His Glu Ala Val Cys Ser Ser			
	300	305	310
gaa cca gaa aat gtt cag gct gac cca gag aat tat acc agc ctt ctt			1134
Glu Pro Glu Asn Val Gln Ala Asp Pro Glu Asn Tyr Thr Ser Leu Leu			
	315	320	325
gtt aca tgg gaa aga cct cga gtc gtt tat gat acc atg att gag aag			1182
Val Thr Trp Glu Arg Pro Arg Val Val Tyr Asp Thr Met Ile Glu Lys			
	330	335	340
ttt gca gtt ttg tac cag cag ttg gat gga gag gac caa acc aag cat			1230
Phe Ala Val Leu Tyr Gln Gln Leu Asp Gly Glu Asp Gln Thr Lys His			
	350	355	360
gaa ttt ttg aca gat ggc tat caa gac ttg gta act ata tga			1272
Glu Phe Leu Thr Asp Gly Tyr Gln Asp Leu Val Thr Ile			
	365	370	
tcagttggtt tacatagggg aacattataa ttttaatttcc aaggtaagaa cttacaaatg			1332
gttggtatatt atttttctcc attactttta gactttatgt gaaggtgggg taggctgagt			1392
atttttaaat ttaaaaaaaaa atttttaaatt agaagctata cttaaattatg ttttaaagtt			1452
catttaatta aatggatata ataactttgc caacaataac actatagagt agatacatat			1512
gacttatgaa ctggagatca tttagtgtgg cttttcttaa gatttcagtt gtagaatagt			1572
gccagaatct cagtgccttg atacatttta tattgtgtct tccattacgc tatatcagca			1632
caggaaaaat agagtgggg acatacaagt cctcttttctt gcacccaaaaa attttcagat			1692

ggtgtagttg acaaaccttc atctggatgt aggtccagaa agtccccact gcagggttaaa	1932
ggacactgga ctctgcactc aggcacctag agtccctgcaa gtccctgggaa cctgcattta	1992
aataaaaaatg cactattaat tatgtttcat atcatgtgga caaaatggat aaaatttttag	2052
taacctttta attcagttgc ctggaatatg gagacacaat gacctgggaa aatcgtgaaa	2112
taaatagtaa taaaaatggt tatttcataa ttacgtgaag aagataattc tattactggt	2172
cttgcatata tattgtcaag aaaaagagat aacttagttg ttactttttt cacattgctc	2232
cttgtttgca aatgcccccc atttatttgt ctaaaatatt aatttttagt ttgtagtact	2292
aatttatgaa tttgatgagt tctggctaaa aatgaaactt cctgaaacta aatctgattt	2352
ttaaaaagca aaaaaaaaaa aaaagcctag ctttccagtt cttcataatt cacaaatacc	2412
acaagtttaa ctaagcaaca ttgcataaac ttttccttag gttaataaaa tagaagtatt	2472
ttccacggac cagggagaaa aagttttcta ggaaagatac ctagtgtggt ggtagtccta	2532
tgagaataac atttgtataa ttactaacat ctttctttta ggggtgctatt ctcaataatt	2592
tgctacccaa tatgagttat gttcttcaga tagtagccat atgcactaat ggcttatatg	2652
gaaaatacag cgaccaactg attgtcgaca tgctactga taatcctggg aagtgccacc	2712
agatacatct atatattaac tcaataaatg aggttagttt aattactgta tgcattgatg	2772
ctttctctct atattctttt ggccaaaagg caaagtgatt ttctcttaag tctggattgc	2832
cgggtaattt tttggggcat gggaccatt tctcattcag caggtctggg gccagacaat	2892
aagtaaactt atccttaata ttggagttta ccatttgtaa aataagagtg actaaacata	2952
tttataacat tgtaataatc attaaatgaa aattgctatg taaatgttga gactgttatt	3012
ttggataatt aagagttggg ttaatttgta tttatttctt cttttcagcc cccaaagcat	3072
tatgtagtaa gtgtataca	3091

<210> 2  
 <211> 374  
 <212> PRT  
 <213> Homo sapiens

2220

24 1

Met Arg Ile Leu Lys Arg Phe Leu Ala Cys Ile Gln Leu Leu Cys Val  
1 5 10 15

Cys Arg Leu Asp Trp Ala Asn Gly Tyr Tyr Arg Gln Gln Arg Lys Leu  
20 25 30

Val Glu Glu Ile Gly Trp Ser Tyr Thr Gly Ala Leu Asn Gln Lys Asn  
35 40 45

Trp Gly Lys Lys Tyr Pro Thr Cys Asn Ser Pro Lys Gln Ser Pro Ile  
50 55 60

Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val Asn Leu Lys Lys Leu  
65 70 75 80

Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu Glu Asn Thr Phe Ile His  
85 90 95

Asn Thr Gly Lys Thr Val Glu Ile Asn Leu Thr Asn Asp Tyr Arg Val  
100 105 110

Ser Gly Gly Val Ser Glu Met Val Phe Lys Ala Ser Lys Ile Thr Phe  
115 120 125

His Trp Gly Lys Cys Asn Met Ser Ser Asp Gly Ser Glu His Ser Leu  
130 135 140

Glu Gly Gln Lys Phe Pro Leu Glu Met Gln Ile Tyr Cys Phe Asp Ala  
145 150 155 160

Asp Arg Phe Ser Ser Phe Glu Glu Ala Val Lys Gly Lys Gly Lys Leu  
165 170 175

Arg Ala Leu Ser Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp  
180 185 190

Phe Lys Ala Ile Ile Asp Gly Val Glu Ser Val Ser Arg Phe Gly Lys  
195 200 205

Thr Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys  
225 230 235 240

Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile Ser  
245 250 255

Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln Gln Ser  
260 265 270

Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe Arg Glu Gln  
275 280 285

Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr Thr Gly Lys Glu  
290 295 300

Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro Glu Asn Val Gln Ala  
305 310 315 320

Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val Thr Trp Glu Arg Pro Arg  
325 330 335

Val Val Tyr Asp Thr Met Ile Glu Lys Phe Ala Val Leu Tyr Gln Gln  
340 345 350

Leu Asp Gly Glu Asp Gln Thr Lys His Glu Phe Leu Thr Asp Gly Tyr  
355 360 365

Gln Asp Leu Val Thr Ile  
370

<210> 3  
<211> 8058  
<212> DNA  
<213> Homo sapiens

<220>  
<221> gene  
<222> (1)..(8058)  
<223> PTP-zeta SM2 exon 23a variant

<224>

..  
<225> .....

&lt;223&gt; 3' Untranslated Region

cacacatacgc	cacgcacgat	ctcacttcga	tctatacact	ggaggattaa	aacaaaacaa	60
caaaaaaaaaac	atttctcttcg	ctccccctcc	ctctccactc	tgagaagcag	aggagccgca	120
cggcgaggggg	ccgcagaccg	tctggaaatg	cgaatcctaa	agcgtttcct	cgtttgcatt	180
cagctcctct	gtgtttgcgc	cctggattgg	gctaattgat	actacagaca	acagagaaaa	240
cttgttgaag	agattggctg	gtcctataca	ggagcactga	atcaaaaaaa	ttggggaaaag	300
aaatatccaa	catgtaatag	cccaaaaaca	tctcctatca	atattgatga	agatcttaca	360
caagtaaatg	tgaatcttaa	gaaacttaaa	tttcaggggt	gggataaaaac	atcattggaa	420
aacacattca	ttcataacac	tgggaaaaca	gtggaaatta	atctcactaa	tgactaccgt	480
gtcagcggag	gagtttcaga	aatggtgttt	aaagcaagca	agataacttt	tactgggga	540
aaatgcaata	tgtcatctga	tggatcagag	catagtttag	aaggacaaaa	atttccactt	600
gagatgcaaa	tctactgctt	tgatgcggac	cgattttcaa	gttttgagga	agcagtcaaa	660
ggaaaaggga	agttaagagc	tttatccatt	ttgtttgagg	ttgggacaga	agaaaatttg	720
gatttcaaag	cgattattga	tggagtcgaa	agtgttagtc	gttttgggaa	gcaggctgct	780
ttagatccat	tcatactgtt	gaacctttctg	ccaaactcaa	ctgacaagta	ttacatttac	840
aatggctcat	tgacatctcc	tccctgcaca	gacacagttg	actggattgt	ttttaagat	900
acagtttagca	tctctgaaag	ccagttgggt	gttttttgtg	aagttcttac	aatgcaacaa	960
tctggttatg	tcatgctgat	ggactactta	caaaacaatt	ttcgagagca	acagtacaag	1020
ttctctagac	aggtgttttc	ctcatacact	ggaaaggaag	agattcatga	agcagtttgt	1080
agttcagaac	cagaaaatgt	tcaggctgac	ccagagaatt	ataccagcct	tcttggtaca	1140
tgggaaagac	ctcgagtcgt	ttatgatacc	atgattgaga	agtttcagct	tttgtaccag	1200
cagttggatg	gagaggacca	aaccaagcat	gaatttttga	cagatggcta	tcaagacttg	1260
ggtgctattc	tcaataattt	gtaccccaat	atgagttatg	ttcttcagat	agtagccata	1320
tgcactaatg	gcttatatgg	aaaatacagc	gaccaactga	ttgtcgacat	gcctactgat	1380





actgaactgc aaattccttc ttccaatgag atgggtttacc cttctgaaag cacagtcatg	3360
cccaacatgt atgataatgt aaataagttg aatgcgtctt tacaagaaac ctctgtttcc	3420
atctctagca ccaagggcat gtttccaggg tcccttgctc ataccaccac taagggtttt	3480
gatcatgaga ttagtcaagt tccagaaaat aacttttcag tccaacctac acatactgtc	3540
tccaagcat ctgggtgacac ttcgcttaaa cctgtgctta gtgcaaacctc agagccagca	3600
tctctgacc ctgcttctag tgaaatgtta tctccttcaa ctgagctctt attttatgag	3660
acctcagctt ctttttagtac tgaagtattg ctacaacctt cctttcaggc ttctgatggt	3720
gacaccttgc ttaaaactgt tcttccagct gtgcccagtg atccaatatt ggttgaaacc	3780
cccaaagttg ataaaattag ttctacaatg ttgcatctca ttgtatcaaa ttctgcttca	3840
agtgaaaaca tgcctgacct tacatctgta ccagtttttg atgtgtcgcc tacttctcat	3900
atgcactctg cttcacttca aggtttgacc atttctatg caagtgagaa atatgaacca	3960
gttttggtta aaagtgaag ttcccaccaa gtggtacctt ctttgtacag taatgatgag	4020
ttgttccaaa cggccaattt ggagattaac caggcccctc ccccaaaagg aaggcatgta	4080
tttgctacac ctgttttctc aattgatgaa ccattaaata cactaataaa taagcttata	4140
cattccgatg aaattttaac ctccaccaa agttctgtta ctggtaaggt atttgcctgt	4200
attccaacag ttgcttctga tacatttgta tctactgac attctgttcc tataggaaat	4260
gggcatgttg ccattacagc tgtttctccc cacagagatg gttctgtaac ctcaacaaag	4320
ttgctgttcc cttctaaggc aacttctgag ctgagtcata gtgcaaaatc tgatgcoggt	4380
ttagtgggtg gtggtgaaga tggtgacact gatgatgatg gtgatgatga tgatgacaga	4440
gatagtgatg gcttatccat tcataagtgt atgtcatgct catcttatag agaatcacag	4500
gaaaaggtaa tgaatgattc agacacccac gaaaacagtc ttatggatca gaataatcga	4560
atctcact cactatctga gaattctgaa gaagataata gagtcacaag tgtatctcga	4620
gacagtcaaa ctgggtatga cagaagtcct ggtaaatcac catcagcaaa tgggctatcc	4680
caaaagcaca atgatggaaa agaggaaaat gacattcaga ctggtagtgc tctgcttctc	4740
ctcagcctg aatctaaagc atgggcagtt ctgacaagtg atgaagaaag tggatcaggg	4800



gaagaacaca aatgtctatc taatgaggaa aaacttataa ttcaggactt tatcttagaa	6780
gctacacagg atgattatgt acttgaagtg aggcactttc agtgtcctaa atggccaaat	6840
ccagatagcc ccattagtaa aacttttgaa cttataagtg ttataaaaga agaagctgcc	6900
aatagggatg ggccatgat tgttcatgat gagcatggag gagtgcggc aggaactttc	6960
tgtgctctga caacccttat gcaccaacta gaaaaagaaa attccgtgga tgtttaccag	7020
gtagccaaga tgatcaatct gatgaggcca ggagtctttg ctgacattga gcagtatcag	7080
tttctctaca aagtgatect cagccttctg agcacaaggc aggaagagaa tccatccacc	7140
tctctggaca gtaatggtgc agcattgect gatggaaata tagctgagag cttagagtct	7200
ttagtttaac acagaaaggg gtggggggac tcacatctga gcattgtttt cctcttcta	7260
aaattaggca ggaaaatcag tctagttctg ttatctgttg atttcccatc acctgacagt	7320
aactttcatg acataggatt ctgcgcgcaa atttatatca ttaacaatgt gtgccttttt	7380
gcaagacttg taatttactt attatgtttg aactaaaatg attgaatttt acagtatttc	7440
taagaatgga attgtggtat tttttctctg attgatttta acagaaaatt tcaatttata	7500
gaggtttaga attccaaact acagaaaatg tttgttttta gtgtcaaatt tttagctgta	7560
tttgtagcaa ttatcaggtt tgctagaaat ataactttta atacagtagc ctgtaaataa	7620
aacactcttc catatgatat tcaacatttt acaactgcag tattcaccta aagtagaaat	7680
aatctgttac ttattgtaaa tactgccta gtgtctccat ggaccaaatt tatatttata	7740
attgtagatt tttatatatt actactgagt caagttttct agttctgtgt aattgtttag	7800
tttaatgacg tagttcatta gctggtctta ctctaccagt tttctgacat tgtattgtgt	7860
tacctaaagt attaaacttg tttcagcatg taattttaac tttgtggaa aatagaaata	7920
ccttcatttt gaaagaagtt tttatgagaa taacacctta ccaaacattg ttcaaatggt	7980
ttttatccaa ggaattgcaa aaataaatat aaatattgcc attaaaaaaaa aaaaaaaaaa	8040
aaaaaaaaaa aaaaaaaaaa	8058

<210> 4  
 <211> 2353  
 <212> PRT

...  
 ...

<223> PTP-zeta SM2 23a exon variant

<400> 4

Met Arg Ile Leu Lys Arg Phe Leu Ala Cys Ile Gln Leu Leu Cys Val  
1 5 10 15

Cys Arg Leu Asp Trp Ala Asn Gly Tyr Tyr Arg Gln Gln Arg Lys Leu  
20 25 30

Val Glu Glu Ile Gly Trp Ser Tyr Thr Gly Ala Leu Asn Gln Lys Asn  
35 40 45

Trp Gly Lys Lys Tyr Pro Thr Cys Asn Ser Pro Lys Gln Ser Pro Ile  
50 55 60

Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val Asn Leu Lys Lys Leu  
65 70 75 80

Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu Glu Asn Thr Phe Ile His  
85 90 95

Asn Thr Gly Lys Thr Val Glu Ile Asn Leu Thr Asn Asp Tyr Arg Val  
100 105 110

Ser Gly Gly Val Ser Glu Met Val Phe Lys Ala Ser Lys Ile Thr Phe  
115 120 125

His Trp Gly Lys Cys Asn Met Ser Ser Asp Gly Ser Glu His Ser Leu  
130 135 140

Glu Gly Gln Lys Phe Pro Leu Glu Met Gln Ile Tyr Cys Phe Asp Ala  
145 150 155 160

Asp Arg Phe Ser Ser Phe Glu Glu Ala Val Lys Gly Lys Gly Lys Leu  
165 170 175

Arg Ala Leu Ser Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp  
180 185 190

Gln Ala Ala Leu Asp Pro Phe Ile Leu Leu Asn Leu Leu Pro Asn Ser  
210 215 220

Thr Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys  
225 230 235 240

Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile Ser  
245 250 255

Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln Gln Ser  
260 265 270

Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe Arg Glu Gln  
275 280 285

Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr Thr Gly Lys Glu  
290 295 300

Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro Glu Asn Val Gln Ala  
305 310 315 320

Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val Thr Trp Glu Arg Pro Arg  
325 330 335

Val Val Tyr Asp Thr Met Ile Glu Lys Phe Ala Val Leu Tyr Gln Gln  
340 345 350

Leu Asp Gly Glu Asp Gln Thr Lys His Glu Phe Leu Thr Asp Gly Tyr  
355 360 365

Gln Asp Leu Gly Ala Ile Leu Asn Asn Leu Leu Pro Asn Met Ser Tyr  
370 375 380

Val Leu Gln Ile Val Ala Ile Cys Thr Asn Gly Leu Tyr Gly Lys Tyr  
385 390 395 400

Ser Asp Gln Leu Ile Val Asp Met Pro Thr Asp Asn Pro Glu Leu Asp  
405 410 415

Leu Tyr Tyr Asp Ile Val His Tyr Ala Ile Val Asn Ile Tyr Asp Asp

435

440

445

Ser Ala Thr Asn Gln Ile Arg Lys Lys Glu Pro Gln Ile Ser Thr Thr  
450 455 460

Thr His Tyr Asn Arg Ile Gly Thr Lys Tyr Asn Glu Ala Lys Thr Asn  
465 470 475 480

Arg Ser Pro Thr Arg Gly Ser Glu Phe Ser Gly Lys Gly Asp Val Pro  
485 490 495

Asn Thr Ser Leu Asn Ser Thr Ser Gln Pro Val Thr Lys Leu Ala Thr  
500 505 510

Glu Lys Asp Ile Ser Leu Thr Ser Gln Thr Val Thr Glu Leu Pro Pro  
515 520 525

His Thr Val Glu Gly Thr Ser Ala Ser Leu Asn Asp Gly Ser Lys Thr  
530 535 540

Val Leu Arg Ser Pro His Met Asn Leu Ser Gly Thr Ala Glu Ser Leu  
545 550 555 560

Asn Thr Val Ser Ile Thr Glu Tyr Glu Glu Glu Ser Leu Leu Thr Ser  
565 570 575

Phe Lys Leu Asp Thr Gly Ala Glu Asp Ser Ser Gly Ser Ser Pro Ala  
580 585 590

Thr Ser Ala Ile Pro Phe Ile Ser Glu Asn Ile Ser Gln Gly Tyr Ile  
595 600 605

Phe Ser Ser Glu Asn Pro Glu Thr Ile Thr Tyr Asp Val Leu Ile Pro  
610 615 620

Glu Ser Ala Arg Asn Ala Ser Glu Asp Ser Thr Ser Ser Gly Ser Glu  
625 630 635 640

Phe Leu Gln Thr Asn Tyr Thr Glu Ile Arg Val Asp Glu Ser Glu Lys  
675 680 685

Thr Thr Lys Ser Phe Ser Ala Gly Pro Val Met Ser Gln Gly Pro Ser  
690 695 700

Val Thr Asp Leu Glu Met Pro His Tyr Ser Thr Phe Ala Tyr Phe Pro  
705 710 715 720

Thr Glu Val Thr Pro His Ala Phe Thr Pro Ser Ser Arg Gln Gln Asp  
725 730 735

Leu Val Ser Thr Val Asn Val Val Tyr Ser Gln Thr Thr Gln Pro Val  
740 745 750

Tyr Asn Gly Glu Thr Pro Leu Gln Pro Ser Tyr Ser Ser Glu Val Phe  
755 760 765

Pro Leu Val Thr Pro Leu Leu Leu Asp Asn Gln Ile Leu Asn Thr Thr  
770 775 780

Pro Ala Ala Ser Ser Ser Asp Ser Ala Leu His Ala Thr Pro Val Phe  
785 790 795 800

Pro Ser Val Asp Val Ser Phe Glu Ser Ile Leu Ser Ser Tyr Asp Gly  
805 810 815

Ala Pro Leu Leu Pro Phe Ser Ser Ala Ser Phe Ser Ser Glu Leu Phe  
820 825 830

Asn His Leu His Thr Val Ser Gln Ile Leu Pro Gln Val Thr Ser Ala  
835 840 845

Thr Glu Ser Asp Lys Val Pro Leu His Ala Ser Leu Pro Val Ala Gly  
850 855 860

Gly Asp Leu Leu Leu Glu Pro Ser Leu Ala Gln Tyr Ser Asp Val Leu

Gly Val Leu Tyr Lys Thr Leu Met Phe Ser Gln Val Glu Pro Pro Ser  
900 905 910

Ser Asp Ala Met Met His Ala Arg Ser Ser Gly Pro Glu Pro Ser Tyr  
915 920 925

Ala Leu Ser Asp Asn Glu Gly Ser Gln His Ile Phe Thr Val Ser Tyr  
930 935 940

Ser Ser Ala Ile Pro Val His Asp Ser Val Gly Val Thr Tyr Gln Gly  
945 950 955 960

Ser Leu Phe Ser Gly Pro Ser His Ile Pro Ile Pro Lys Ser Ser Leu  
965 970 975

Ile Thr Pro Thr Ala Ser Leu Leu Gln Pro Thr His Ala Leu Ser Gly  
980 985 990

Asp Gly Glu Trp Ser Gly Ala Ser Ser Asp Ser Glu Phe Leu Leu Pro  
995 1000 1005

Asp Thr Asp Gly Leu Thr Ala Leu Asn Ile Ser Ser Pro Val Ser  
1010 1015 1020

Val Ala Glu Phe Thr Tyr Thr Thr Ser Val Phe Gly Asp Asp Asn  
1025 1030 1035

Lys Ala Leu Ser Lys Ser Glu Ile Ile Tyr Gly Asn Glu Thr Glu  
1040 1045 1050

Leu Gln Ile Pro Ser Phe Asn Glu Met Val Tyr Pro Ser Glu Ser  
1055 1060 1065

Thr Val Met Pro Asn Met Tyr Asp Asn Val Asn Lys Leu Asn Ala  
1070 1075 1080

Ser Leu Gln Glu Thr Ser Val Ser Ile Ser Ser Thr Lys Gly Met  
1085 1090 1095



Glu Ile Ser Gln Val Pro Glu Asn Asn Phe Ser Val Gln Pro Thr  
1115 1120 1125

His Thr Val Ser Gln Ala Ser Gly Asp Thr Ser Leu Lys Pro Val  
1130 1135 1140

Leu Ser Ala Asn Ser Glu Pro Ala Ser Ser Asp Pro Ala Ser Ser  
1145 1150 1155

Glu Met Leu Ser Pro Ser Thr Gln Leu Leu Phe Tyr Glu Thr Ser  
1160 1165 1170

Ala Ser Phe Ser Thr Glu Val Leu Leu Gln Pro Ser Phe Gln Ala  
1175 1180 1185

Ser Asp Val Asp Thr Leu Leu Lys Thr Val Leu Pro Ala Val Pro  
1190 1195 1200

Ser Asp Pro Ile Leu Val Glu Thr Pro Lys Val Asp Lys Ile Ser  
1205 1210 1215

Ser Thr Met Leu His Leu Ile Val Ser Asn Ser Ala Ser Ser Glu  
1220 1225 1230

Asn Met Leu His Ser Thr Ser Val Pro Val Phe Asp Val Ser Pro  
1235 1240 1245

Thr Ser His Met His Ser Ala Ser Leu Gln Gly Leu Thr Ile Ser  
1250 1255 1260

Tyr Ala Ser Glu Lys Tyr Glu Pro Val Leu Leu Lys Ser Glu Ser  
1265 1270 1275

Ser His Gln Val Val Pro Ser Leu Tyr Ser Asn Asp Glu Leu Phe  
1280 1285 1290

Gln Thr Ala Asn Leu Glu Ile Asn Gln Ala His Pro Pro Lys Gly  
1295 1300 1305

Asn Thr Val Leu Asn Lys Leu Ile His Ser Asp Val Leu Leu Thr

1325		1330		1335
Ser Thr Lys Ser Ser Val Thr Gly Lys Val Phe Ala Gly Ile Pro				
1340		1345		1350
Thr Val Ala Ser Asp Thr Phe Val Ser Thr Asp His Ser Val Pro				
1355		1360		1365
Ile Gly Asn Gly His Val Ala Ile Thr Ala Val Ser Pro His Arg				
1370		1375		1380
Asp Gly Ser Val Thr Ser Thr Lys Leu Leu Phe Pro Ser Lys Ala				
1385		1390		1395
Thr Ser Glu Leu Ser His Ser Ala Lys Ser Asp Ala Gly Leu Val				
1400		1405		1410
Gly Gly Gly Glu Asp Gly Asp Thr Asp Asp Asp Gly Asp Asp Asp				
1415		1420		1425
Asp Asp Arg Asp Ser Asp Gly Leu Ser Ile His Lys Cys Met Ser				
1430		1435		1440
Cys Ser Ser Tyr Arg Glu Ser Gln Glu Lys Val Met Asn Asp Ser				
1445		1450		1455
Asp Thr His Glu Asn Ser Leu Met Asp Gln Asn Asn Pro Ile Ser				
1460		1465		1470
Tyr Ser Leu Ser Glu Asn Ser Glu Glu Asp Asn Arg Val Thr Ser				
1475		1480		1485
Val Ser Ser Asp Ser Gln Thr Gly Met Asp Arg Ser Pro Gly Lys				
1490		1495		1500
Ser Pro Ser Ala Asn Gly Leu Ser Gln Lys His Asn Asp Gly Lys				
1505		1510		1515

Gly Ser Gly Gln Gly Thr Ser Asp Ser Leu Asn Glu Asn Glu Thr  
1550 1555 1560

Ser Thr Asp Phe Ser Phe Ala Asp Thr Asn Glu Lys Asp Ala Asp  
1565 1570 1575

Gly Ile Leu Ala Ala Gly Asp Ser Glu Ile Thr Pro Gly Phe Pro  
1580 1585 1590

Gln Ser Pro Thr Ser Ser Val Thr Ser Glu Asn Ser Glu Val Phe  
1595 1600 1605

His Val Ser Glu Ala Glu Ala Ser Asn Ser Ser His Glu Ser Arg  
1610 1615 1620

Ile Gly Leu Ala Glu Gly Leu Glu Ser Glu Lys Lys Ala Val Ile  
1625 1630 1635

Pro Leu Val Ile Val Ser Ala Leu Thr Phe Ile Cys Leu Val Val  
1640 1645 1650

Leu Val Gly Ile Leu Ile Tyr Trp Arg Lys Cys Phe Gln Thr Ala  
1655 1660 1665

His Phe Tyr Leu Glu Asp Ser Thr Ser Pro Arg Val Ile Ser Thr  
1670 1675 1680

Pro Pro Thr Pro Ile Phe Pro Ile Ser Asp Asp Val Gly Ala Ile  
1685 1690 1695

Pro Ile Lys His Phe Pro Lys His Val Ala Asp Leu His Ala Ser  
1700 1705 1710

Ser Gly Phe Thr Glu Glu Phe Glu Thr Leu Lys Glu Phe Tyr Gln  
1715 1720 1725

Glu Val Glu Ser Cys Thr Val Asp Leu Gly Ile Thr Ala Asp Ser

Val Ala Tyr Asp His Ser Arg Val Lys Leu Ala Gln Leu Ala Glu  
1760 1765 1770

Lys Asp Gly Lys Leu Thr Asp Tyr Ile Asn Ala Asn Tyr Val Asp  
1775 1780 1785

Gly Tyr Asn Arg Pro Lys Ala Tyr Ile Ala Ala Gln Gly Pro Leu  
1790 1795 1800

Lys Ser Thr Ala Glu Asp Phe Trp Arg Met Ile Trp Glu His Asn  
1805 1810 1815

Val Glu Val Ile Val Met Ile Thr Asn Leu Val Glu Lys Gly Arg  
1820 1825 1830

Arg Lys Cys Asp Gln Tyr Trp Pro Ala Asp Gly Ser Glu Glu Tyr  
1835 1840 1845

Gly Asn Phe Leu Val Thr Gln Lys Ser Val Gln Val Leu Ala Tyr  
1850 1855 1860

Tyr Thr Val Arg Asn Phe Thr Leu Arg Asn Thr Lys Ile Lys Lys  
1865 1870 1875

Gly Ser Gln Lys Gly Arg Pro Ser Gly Arg Val Val Thr Gln Tyr  
1880 1885 1890

His Tyr Thr Gln Trp Pro Asp Met Gly Val Pro Glu Tyr Ser Leu  
1895 1900 1905

Pro Val Leu Thr Phe Val Arg Lys Ala Ala Tyr Ala Lys Arg His  
1910 1915 1920

Ala Val Gly Pro Val Val Val His Cys Ser Ala Gly Val Gly Arg  
1925 1930 1935

Thr Gly Thr Tyr Ile Val Leu Asp Ser Met Leu Gln Gln Ile Gln  
1940 1945 1950

Ser Gln Arg Asn Tyr Leu Val Gln Thr Glu Glu Gln Tyr Val Phe  
1970 1975 1980

Ile His Asp Thr Leu Val Glu Ala Ile Leu Ser Lys Glu Thr Glu  
1985 1990 1995

Val Leu Asp Ser His Ile His Ala Tyr Val Asn Ala Leu Leu Ile  
2000 2005 2010

Pro Gly Pro Ala Gly Lys Thr Lys Leu Glu Lys Gln Phe Gln Gly  
2015 2020 2025

Leu Thr Leu Ser Pro Arg Leu Glu Cys Arg Gly Thr Ile Ser Ala  
2030 2035 2040

His Cys Asn Leu Pro Leu Pro Gly Leu Thr Asp Pro Pro Thr Ser  
2045 2050 2055

Ala Ser Arg Val Ala Gly Thr Ile Leu Leu Ser Gln Ser Asn Ile  
2060 2065 2070

Gln Gln Ser Asp Tyr Ser Ala Ala Leu Lys Gln Cys Asn Arg Glu  
2075 2080 2085

Lys Asn Arg Thr Ser Ser Ile Ile Pro Val Glu Arg Ser Arg Val  
2090 2095 2100

Gly Ile Ser Ser Leu Ser Gly Glu Gly Thr Asp Tyr Ile Asn Ala  
2105 2110 2115

Ser Tyr Ile Met Gly Tyr Tyr Gln Ser Asn Glu Phe Ile Ile Thr  
2120 2125 2130

Gln His Pro Leu Leu His Thr Ile Lys Asp Phe Trp Arg Met Ile  
2135 2140 2145

Trp Asp His Asn Ala Gln Leu Val Val Met Ile Pro Asp Gly Gln  
2150 2155 2160

2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 9835 9840 9845 9850 9855 9860 9865 9870 9875 9880 9885 9890 9895 9900 9905 9910 9915 9920 9925 9930 9935 9940 9945 9950 9955 9960 9965 9970 9975 9980 9985 9990 9995 10000

2180	2185	2190
His Lys Cys Leu Ser Asn Glu Glu Lys Leu Ile Ile Gln Asp Phe		
2195	2200	2205
Ile Leu Glu Ala Thr Gln Asp Asp Tyr Val Leu Glu Val Arg His		
2210	2215	2220
Phe Gln Cys Pro Lys Trp Pro Asn Pro Asp Ser Pro Ile Ser Lys		
2225	2230	2235
Thr Phe Glu Leu Ile Ser Val Ile Lys Glu Glu Ala Ala Asn Arg		
2240	2245	2250
Asp Gly Pro Met Ile Val His Asp Glu His Gly Gly Val Thr Ala		
2255	2260	2265
Gly Thr Phe Cys Ala Leu Thr Thr Leu Met His Gln Leu Glu Lys		
2270	2275	2280
Glu Asn Ser Val Asp Val Tyr Gln Val Ala Lys Met Ile Asn Leu		
2285	2290	2295
Met Arg Pro Gly Val Phe Ala Asp Ile Glu Gln Tyr Gln Phe Leu		
2300	2305	2310
Tyr Lys Val Ile Leu Ser Leu Val Ser Thr Arg Gln Glu Glu Asn		
2315	2320	2325
Pro Ser Thr Ser Leu Asp Ser Asn Gly Ala Ala Leu Pro Asp Gly		
2330	2335	2340
Asn Ile Ala Glu Ser Leu Glu Ser Leu Val		
2345	2350	

<210> 5  
 <211> 7941  
 <212> DNA  
 <213> Homo sapiens

<220>

<221> gene

<222> (1)..(7941)

<223> PTP-zeta

<400> 5

cacacataag cagcagcat ctcaattaga tctatacact ggaggattaa aacaaacaaa 60

caaaaaaaaaac atttctttcg ctccccctcc ctctccactc tgagaagcag aggagcgcga 120

cggcgagggg ccgcagaccg tctggaa atg cga atc cta aag cgt ttc ctc gct 174

Met Arg Ile Leu Lys Arg Phe Leu Ala

1

5

tgc att cag ctc ctc tgt gtt tgc cgc ctg gat tgg gct aat gga tac 222

Cys Ile Gln Leu Leu Cys Val Cys Arg Leu Asp Trp Ala Asn Gly Tyr

10

15

20

25

tac aga caa cag aga aaa ctt gtt gaa gag att ggc tgg tcc tat aca 270

Tyr Arg Gln Gln Arg Lys Leu Val Glu Glu Ile Gly Trp Ser Tyr Thr

30

35

40

gga gca ctg aat caa aaa aat tgg gga aag aaa tat cca aca tgt aat 318

Gly Ala Leu Asn Gln Lys Asn Trp Gly Lys Lys Tyr Pro Thr Cys Asn

45

50

55

agc cca aaa caa tct cct atc aat att gat gaa gat ctt aca caa gta 366

Ser Pro Lys Lys Gln Ser Pro Ile Asn Ile Asp Glu Asp Leu Thr Gln Val

60

65

70

aat gtg aat ctt aag aaa ctt aaa ttt cag ggt tgg gat aaa aca tca 414

Asn Val Asn Leu Lys Lys Leu Lys Phe Gln Gly Trp Asp Lys Thr Ser

75

80

85

ttg gaa aac aca ttc att cat aac act ggg aaa aca gtg gaa att aat 462

Leu Glu Asn Thr Phe Ile His Asn Thr Gly Lys Thr Val Glu Ile Asn

90

95

100

105

ctc act aat gac tac cgt gtc agc gga gga gtt tca gaa atg gtg ttt 510

Leu Thr Asn Asp Tyr Arg Val Ser Gly Gly Val Ser Glu Met Val Phe

110

115

120

aaa gca agc aag ata act ttt cac tgg gga aaa tgc aat atg tca tct 558

Lys Ala Ser Lys Ile Thr Phe His Trp Gly Lys Cys Asn Met Ser Ser

125

130

135

gat gga tca gag cat agt tta gaa gga caa aaa ttt cca ctt gag atg 606

Asp Gly Ser Glu His Ser Leu Glu Gly Gln Lys Phe Pro Leu Glu Met

140

145

150

155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995

Val	Lys	Gly	Lys	Gly	Lys	Leu	Arg	Ala	Leu	Ser	Ile	Leu	Phe	Glu	Val	
170					175					180					185	
ggg	aca	gaa	gaa	aat	ttg	gat	ttc	aaa	gcg	att	att	gat	gga	gtc	gaa	750
Gly	Thr	Glu	Glu	Asn	Leu	Asp	Phe	Lys	Ala	Ile	Ile	Asp	Gly	Val	Glu	
				190				195					200			
agt	gtt	agt	cgt	ttt	ggg	aag	cag	gct	gct	tta	gat	cca	ttc	ata	ctg	798
Ser	Val	Ser	Arg	Phe	Gly	Lys	Gln	Ala	Ala	Leu	Asp	Pro	Phe	Ile	Leu	
			205					210					215			
ttg	aac	ctt	ctg	cca	aac	tca	act	gac	aag	tat	tac	att	tac	aat	ggc	846
Leu	Asn	Leu	Leu	Pro	Asn	Ser	Thr	Asp	Lys	Tyr	Tyr	Ile	Tyr	Asn	Gly	
		220					225					230				
tca	ttg	aca	tct	cct	ccc	tgc	aca	gac	aca	gtt	gac	tgg	att	gtt	ttt	894
Ser	Leu	Thr	Ser	Pro	Pro	Cys	Thr	Asp	Thr	Val	Asp	Trp	Ile	Val	Phe	
	235					240					245					
aaa	gat	aca	gtt	agc	atc	tct	gaa	agc	cag	ttg	gct	gtt	ttt	tgt	gaa	942
Lys	Asp	Thr	Val	Ser	Ile	Ser	Glu	Ser	Gln	Leu	Ala	Val	Phe	Cys	Glu	
250					255				260					265		
gtt	ctt	aca	atg	caa	caa	tct	ggg	tat	gtc	atg	ctg	atg	gac	tac	tta	990
Val	Leu	Thr	Met	Gln	Gln	Ser	Gly	Tyr	Val	Met	Leu	Met	Asp	Tyr	Leu	
				270				275						280		
caa	aac	aat	ttt	cga	gag	caa	cag	tac	aag	ttc	tct	aga	cag	gtg	ttt	1038
Gln	Asn	Asn	Phe	Arg	Glu	Gln	Gln	Tyr	Lys	Phe	Ser	Arg	Gln	Val	Phe	
			285					290					295			
tcc	tca	tac	act	gga	aag	gaa	gag	att	cat	gaa	gca	gtt	tgt	agt	tca	1086
Ser	Ser	Tyr	Thr	Gly	Lys	Glu	Glu	Ile	His	Glu	Ala	Val	Cys	Ser	Ser	
		300					305					310				
gaa	cca	gaa	aat	gtt	cag	gct	gac	cca	gag	aat	tat	acc	agc	ctt	ctt	1134
Glu	Pro	Glu	Asn	Val	Gln	Ala	Asp	Pro	Glu	Asn	Tyr	Thr	Ser	Leu	Leu	
	315					320					325					
gtt	aca	tgg	gaa	aga	cct	cga	gtc	gtt	tat	gat	acc	atg	att	gag	aag	1182
Val	Thr	Trp	Glu	Arg	Pro	Arg	Val	Val	Tyr	Asp	Thr	Met	Ile	Glu	Lys	
330					335				340					345		
ttt	gca	gtt	ttg	tac	cag	cag	ttg	gat	gga	gag	gac	caa	acc	aag	cat	1230
Phe	Ala	Val	Leu	Tyr	Gln	Gln	Leu	Asp	Gly	Glu	Asp	Gln	Thr	Lys	His	
			350					355					360			
gaa	ttt	ttg	aca	gat	ggc	tat	caa	gac	ttg	ggg	gct	att	ctc	aat	aat	1278
Glu	Phe	Leu	Thr	Asp	Gly	Tyr	Gln	Asp	Leu	Gly	Ala	Ile	Leu	Asn	Asn	
			365					370					375			

118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135  
 Asn Gly Leu Tyr Gly Lys Tyr Ser Arg Gln Leu Ile Val Arg Met Phe





tca act tca tca ggt tca gaa gaa tca cta aag gat cct tct atg gag Ser Thr Ser Ser Gly Ser Glu Glu Ser Leu Lys Asp Pro Ser Met Glu 635 640 645	2094
gga aat gtg tgg ttt cct agc tct aca gac ata aca gca cag ccc gat Gly Asn Val Trp Phe Pro Ser Ser Thr Asp Ile Thr Ala Gln Pro Asp 650 655 660 665	2142
gtt gga tca ggc aga gag agc ttt ctc cag act aat tac act gag ata Val Gly Ser Gly Arg Glu Ser Phe Leu Gln Thr Asn Tyr Thr Glu Ile 670 675 680	2190
cgt gtt gat gaa tct gag aag aca acc aag tcc ttt tct gca ggc cca Arg Val Asp Glu Ser Glu Lys Thr Thr Lys Ser Phe Ser Ala Gly Pro 685 690 695	2238
gtg atg tca cag ggt ccc tca gtt aca gat ctg gaa atg cca cat tat Val Met Ser Gln Gly Pro Ser Val Thr Asp Leu Glu Met Pro His Tyr 700 705 710	2286
tct acc ttt gcc tac ttc cca act gag gta aca cct cat gct ttt acc Ser Thr Phe Ala Tyr Phe Pro Thr Glu Val Thr Pro His Ala Phe Thr 715 720 725	2334
cca tcc tcc aga caa cag gat ttg gtc tcc acg gtc aac gtg gta tac Pro Ser Ser Arg Gln Gln Asp Leu Val Ser Thr Val Asn Val Val Tyr 730 735 740 745	2382
tcg cag aca acc caa ccg gta tac aat ggt gag aca cct ctt caa cct Ser Gln Thr Thr Gln Pro Val Tyr Asn Gly Glu Thr Pro Leu Gln Pro 750 755 760	2430
tcc tac agt agt gaa gtc ttt cct cta gtc acc cct ttg ttg ctt gac Ser Tyr Ser Ser Glu Val Phe Pro Leu Val Thr Pro Leu Leu Leu Asp 765 770 775	2478
aat cag atc ctc aac act acc cct gct gct tca agt agt gat tcg gcc Asn Gln Ile Leu Asn Thr Thr Pro Ala Ala Ser Ser Ser Asp Ser Ala 780 785 790	2526
ttg cat gct acg cct gta ttt ccc agt gtc gat gtg tca ttt gaa tcc Leu His Ala Thr Pro Val Phe Pro Ser Val Asp Val Ser Phe Glu Ser 795 800 805	2574
atc ctg tct tcc tat gat ggt gca cct ttg ctt cca ttt tcc tct gct Ile Leu Ser Ser Tyr Asp Gly Ala Pro Leu Leu Pro Phe Ser Ser Ala 810 815 820 825	2622
tcc ttc agt agt gaa ttg ttt cgc cat ctg cat aca gtt tct caa atc Ser Phe Ser Ser Glu Leu Phe Arg His Leu His Thr Val Ser Gln Ile	2670

gct tct ctg cca gtg gct ggg ggt gat ttg cta tta gag ccc agc ctt Ala Ser Leu Pro Val Ala Gly Gly Asp Leu Leu Leu Glu Pro Ser Leu 860 865 870	2765
gct cag tat tct gat gtg ctg tcc act act cat gct gct tca gag acg Ala Gln Tyr Ser Asp Val Leu Ser Thr Thr His Ala Ala Ser Glu Thr 875 880 885	2814
ctg gaa ttt ggt agt gaa tct ggt gtt ctt tat aaa acg ctt atg ttt Leu Glu Phe Gly Ser Glu Ser Gly Val Leu Tyr Lys Thr Leu Met Phe 890 895 900 905	2862
tct caa gtt gaa cca ccc agc agt gat gcc atg atg cat gca cgt tct Ser Gln Val Glu Pro Pro Ser Ser Asp Ala Met Met His Ala Arg Ser 910 915 920	2910
tca ggg cct gaa cct tct tat gcc ttg tct gat aat gag ggc tcc caa Ser Gly Pro Glu Pro Ser Tyr Ala Leu Ser Asp Asn Glu Gly Ser Gln 925 930 935	2958
cac atc ttc act gtt tct tac agt tct gca ata cct gtg cat gat tct His Ile Phe Thr Val Ser Tyr Ser Ser Ala Ile Pro Val His Asp Ser 940 945 950	3006
gtg ggt gta act tat cag ggt tcc tta ttt agc ggc cct agc cat ata Val Gly Val Thr Tyr Gln Gly Ser Leu Phe Ser Gly Pro Ser His Ile 955 960 965	3054
cca ata cct aag tct tcg tta ata acc cca act gca tca tta ctg cag Pro Ile Pro Lys Ser Ser Leu Ile Thr Pro Thr Ala Ser Leu Leu Gln 970 975 980 985	3102
cct act cat gcc ctg tct ggt gat ggg gaa tgg tct gga gcc tct tct Pro Thr His Ala Leu Ser Gly Asp Gly Glu Trp Ser Gly Ala Ser Ser 990 995 1000	3150
gat agt gaa ttt ctt tta cct gac aca gat ggg ctg aca gcc ctt Asp Ser Glu Phe Leu Leu Pro Asp Thr Asp Gly Leu Thr Ala Leu 1005 1010 1015	3195
aac att tct tca cct gtt tct gta gct gaa ttt aca tat aca aca Asn Ile Ser Ser Pro Val Ser Val Ala Glu Phe Thr Tyr Thr Thr 1020 1025 1030	3240
tct gtg ttt ggt gat gat aat aag gcg ctt tct aaa agt gaa ata Ser Val Phe Gly Asp Asp Asn Lys Ala Leu Ser Lys Ser Glu Ile 1035 1040 1045	3285
ata tat gga aat gag act gaa ctg caa att cct tct ttc aat gag Ile Tyr Gly Asn Glu Thr Glu Leu Gln Ile Pro Ser Phe Asn Glu 1050 1055 1060	3330

atg ata tat gag ttt aat ggt tct tta gaa aat agt tct gtt tct 341

Asn Val Asn Lys	Leu Asn Ala Ser	Leu Gln Glu Thr Ser Val Ser	
1080		1085	1090
att tct agc acc	aag ggc atg ttt cca	ggg tcc ctt gct cat acc	3465
Ile Ser Ser Thr	Lys Gly Met Phe Pro	Gly Ser Leu Ala His Thr	
1095	1100	1105	
acc act aag gtt	ttt gat cat gag att	agt caa gtt cca gaa aat	3510
Thr Thr Lys Val	Phe Asp His Glu Ile	Ser Gln Val Pro Glu Asn	
1110	1115	1120	
aac ttt tca gtt	caa cct aca cat act	gtc tct caa gca tct ggt	3555
Asn Phe Ser Val	Gln Pro Thr His Thr	Val Ser Gln Ala Ser Gly	
1125	1130	1135	
gac act tcg ctt	aaa cct gtg ctt agt	gca aac tca gag cca gca	3600
Asp Thr Ser Leu	Lys Pro Val Leu Ser	Ala Asn Ser Glu Pro Ala	
1140	1145	1150	
tcc tct gac cct	gct tct agt gaa atg	tta tct cct tca act cag	3645
Ser Ser Asp Pro	Ala Ser Ser Glu Met	Leu Ser Pro Ser Thr Gln	
1155	1160	1165	
ctc tta ttt tat	gag acc tca gct tct	ttt agt act gaa gta ttg	3690
Leu Leu Phe Tyr	Glu Thr Ser Ala Ser	Phe Ser Thr Glu Val Leu	
1170	1175	1180	
cta caa cct tcc	ttt cag gct tct gat	gtt gac acc ttg ctt aaa	3735
Leu Gln Pro Ser	Phe Gln Ala Ser Asp	Val Asp Thr Leu Leu Lys	
1185	1190	1195	
act gtt ctt cca	gct gtg ccc agt gat	cca ata ttg gtt gaa acc	3780
Thr Val Leu Pro	Ala Val Pro Ser Asp	Pro Ile Leu Val Glu Thr	
1200	1205	1210	
ccc aaa gtt gat	aaa att agt tct aca	atg ttg cat ctc att gta	3825
Pro Lys Val Asp	Lys Ile Ser Ser Thr	Met Leu His Leu Ile Val	
1215	1220	1225	
tca aat tct gct	tca agt gaa aac atg	ctg cac tct aca tct gta	3870
Ser Asn Ser Ala	Ser Ser Glu Asn Met	Leu His Ser Thr Ser Val	
1230	1235	1240	
cca gtt ttt gat	gtg tcg cct act tct	cat atg cac tct gct tca	3915
Pro Val Phe Asp	Val Ser Pro Thr Ser	His Met His Ser Ala Ser	
1245	1250	1255	
ctt caa ggt ttg	acc att tcc tat gca	agt gag aaa tat gaa cca	3960
Leu Gln Gly Leu	Thr Ile Ser Tyr Ala	Ser Glu Lys Tyr Glu Pro	
1260	1265	1270	

Tyr Ser Asn Asp	Glu Leu Ile Gln Thr	Ala Asn Leu Glu Ile	Asn
-----------------	---------------------	---------------------	-----

1290	1295	1300	
cag gcc cat ccc cca aaa gga agg cat gta ttt gct aca cct gtt			4095
Gln Ala His Pro Pro Lys Gly Arg His Val Phe Ala Thr Pro Val			
1305	1310	1315	
tta tca att gat gaa cca tta aat aca cta ata aat aag ctt ata			4140
Leu Ser Ile Asp Glu Pro Leu Asn Thr Leu Ile Asn Lys Leu Ile			
1320	1325	1330	
cat tcc gat gaa att tta acc tcc acc aaa agt tct gtt act ggt			4185
His Ser Asp Glu Ile Leu Thr Ser Thr Lys Ser Ser Val Thr Gly			
1335	1340	1345	
aag gta ttt gct ggt att cca aca gtt gct tct gat aca ttt gta			4230
Lys Val Phe Ala Gly Ile Pro Thr Val Ala Ser Asp Thr Phe Val			
1350	1355	1360	
tct act gat cat tct gtt cct ata gga aat ggg cat gtt gcc att			4275
Ser Thr Asp His Ser Val Pro Ile Gly Asn Gly His Val Ala Ile			
1365	1370	1375	
aca gct gtt tct ccc cac aga gat ggt tct gta acc tca aca aag			4320
Thr Ala Val Ser Pro His Arg Asp Gly Ser Val Thr Ser Thr Lys			
1380	1385	1390	
ttg ctg ttt cct tct aag gca act tct gag ctg agt cat agt gcc			4365
Leu Leu Phe Pro Ser Lys Ala Thr Ser Glu Leu Ser His Ser Ala			
1395	1400	1405	
aaa tct gat gcc ggt tta gtg ggt ggt ggt gaa gat ggt gac act			4410
Lys Ser Asp Ala Gly Leu Val Gly Gly Gly Glu Asp Gly Asp Thr			
1410	1415	1420	
gat gat gat ggt gat gat gat gat gac aga gat agt gat ggc tta			4455
Asp Asp Asp Gly Asp Asp Asp Asp Asp Arg Asp Ser Asp Gly Leu			
1425	1430	1435	
tcc att cat aag tgt atg tca tgc tca tcc tat aga gaa tca cag			4500
Ser Ile His Lys Cys Met Ser Cys Ser Ser Tyr Arg Glu Ser Gln			
1440	1445	1450	
gaa aag gta atg aat gat tca gac acc cac gaa aac agt ctt atg			4545
Glu Lys Val Met Asn Asp Ser Asp Thr His Glu Asn Ser Leu Met			
1455	1460	1465	
gat cag aat aat cca atc tca tac tca cta tct gag aat tct gaa			4590
Asp Gln Asn Asn Pro Ile Ser Tyr Ser Leu Ser Glu Asn Ser Glu			
1470	1475	1480	
gaa gat aat aga gtc aca agt gta tcc tca gac agt caa act ggt			4635

caa aag cac aat gat gga aaa gag gaa aat gac att cag act ggt	4725
Gln Lys His Asn Asp Gly Lys Glu Glu Asn Asp Ile Gln Thr Gly	
1515 1520 1525	
agt gct ctg ctt cct ctc agc cct gaa tct aaa gca tgg gca gtt	4770
Ser Ala Leu Leu Pro Leu Ser Pro Glu Ser Lys Ala Trp Ala Val	
1530 1535 1540	
ctg aca agt gat gaa gaa agt gga tca ggg caa ggt acc tca gat	4815
Leu Thr Ser Asp Glu Glu Ser Gly Ser Gly Gln Gly Thr Ser Asp	
1545 1550 1555	
agc ctt aat gag aat gag act tcc aca gat ttc agt ttt gca gac	4860
Ser Leu Asn Glu Asn Glu Thr Ser Thr Asp Phe Ser Phe Ala Asp	
1560 1565 1570	
act aat gaa aaa gat gct gat ggg atc ctg gca gca ggt gac tca	4905
Thr Asn Glu Lys Asp Ala Asp Gly Ile Leu Ala Ala Gly Asp Ser	
1575 1580 1585	
gaa ata act cct gga ttc cca cag tcc cca aca tca tct gtt act	4950
Glu Ile Thr Pro Gly Phe Pro Gln Ser Pro Thr Ser Ser Val Thr	
1590 1595 1600	
agc gag aac tca gaa gtg ttc cac gtt tca gag gca gag gcc agt	4995
Ser Glu Asn Ser Glu Val Phe His Val Ser Glu Ala Glu Ala Ser	
1605 1610 1615	
aat agt agc cat gag tct cgt att ggt cta gct gag ggg ttg gaa	5040
Asn Ser Ser His Glu Ser Arg Ile Gly Leu Ala Glu Gly Leu Glu	
1620 1625 1630	
tcc gag aag aag gca gtt ata ccc ctt gtg atc gtg tca gcc ctg	5085
Ser Glu Lys Lys Ala Val Ile Pro Leu Val Ile Val Ser Ala Leu	
1635 1640 1645	
act ttt atc tgt cta gtg gtt ctt gtg ggt att ctc atc tac tgg	5130
Thr Phe Ile Cys Leu Val Val Leu Val Gly Ile Leu Ile Tyr Trp	
1650 1655 1660	
agg aaa tgc ttc cag act gca cac ttt tac tta gag gat agt aca	5175
Arg Lys Cys Phe Gln Thr Ala His Phe Tyr Leu Glu Asp Ser Thr	
1665 1670 1675	
tcc cct aga gtt ata tcc aca cct cca aca cct atc ttt cca att	5220
Ser Pro Arg Val Ile Ser Thr Pro Pro Thr Pro Ile Phe Pro Ile	
1680 1685 1690	
tca gat gat gtc gga gca att cca ata aag cac ttt cca aag cat	5265
Ser Asp Asp Val Gly Ala Ile Pro Ile Lys His Phe Pro Lys His	

aca ctg aaa gag ttt tac cag gaa gtg cag agc tgt act gtt gac	5355
Thr Leu Lys Glu Phe Tyr Gln Glu Val Gln Ser Cys Thr Val Asp	
1725 1730 1735	
tta ggt att aca gca gac agc tcc aac cac cca gac aac aag cac	5400
Leu Gly Ile Thr Ala Asp Ser Ser Asn His Pro Asp Asn Lys His	
1740 1745 1750	
aag aat cga tac ata aat atc gtt gcc tat gat cat agc agg gtt	5445
Lys Asn Arg Tyr Ile Asn Ile Val Ala Tyr Asp His Ser Arg Val	
1755 1760 1765	
aag cta gca cag ctt gct gaa aag gat ggc aaa ctg act gat tat	5490
Lys Leu Ala Gln Leu Ala Glu Lys Asp Gly Lys Leu Thr Asp Tyr	
1770 1775 1780	
atc aat gcc aat tat gtt gat ggc tac aac aga cca aaa gct tat	5535
Ile Asn Ala Asn Tyr Val Asp Gly Tyr Asn Arg Pro Lys Ala Tyr	
1785 1790 1795	
att gct gcc caa ggc cca ctg aaa tcc aca gct gaa gat ttc tgg	5580
Ile Ala Ala Gln Gly Pro Leu Lys Ser Thr Ala Glu Asp Phe Trp	
1800 1805 1810	
aga atg ata tgg gaa cat aat gtg gaa gtt att gtc atg ata aca	5625
Arg Met Ile Trp Glu His Asn Val Glu Val Ile Val Met Ile Thr	
1815 1820 1825	
aac ctc gtg gag aaa gga agg aga aaa tgt gat cag tac tgg cct	5670
Asn Leu Val Glu Lys Gly Arg Arg Lys Cys Asp Gln Tyr Trp Pro	
1830 1835 1840	
gcc gat ggg agt gag gag tac ggg aac ttt ctg gtc act cag aag	5715
Ala Asp Gly Ser Glu Glu Tyr Gly Asn Phe Leu Val Thr Gln Lys	
1845 1850 1855	
agt gtg caa gtg ctt gcc tat tat act gtg agg aat ttt act cta	5760
Ser Val Gln Val Leu Ala Tyr Tyr Thr Val Arg Asn Phe Thr Leu	
1860 1865 1870	
aga aac aca aaa ata aaa aag ggc tcc cag aaa gga aga ccc agt	5805
Arg Asn Thr Lys Ile Lys Lys Gly Ser Gln Lys Gly Arg Pro Ser	
1875 1880 1885	
gga cgt gtg gtc aca cag tat cac tac acg cag tgg cct gac atg	5850
Gly Arg Val Val Thr Gln Tyr His Tyr Thr Gln Trp Pro Asp Met	
1890 1895 1900	
gga gta cca gag tac tcc ctg cca gtg ctg acc ttt gtg aga aag	5895
Gly Val Pro Glu Tyr Ser Leu Pro Val Leu Thr Phe Val Arg Lys	
1905 1910 1915	

1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000

Cys	Ser	Ala	Gly	Val	Gly	Arg	Thr	Gly	Thr	Tyr	Ile	Val	Leu	Asp	
			1935												
								1940					1945		
agt	atg	ttg	cag	cag	att	caa	cac	gaa	gga	act	gtc	aac	ata	ttt	6030
Ser	Met	Leu	Gln	Gln	Ile	Gln	His	Glu	Gly	Thr	Val	Asn	Ile	Phe	
			1950					1955					1960		
ggc	ttc	tta	aaa	cac	atc	cgt	tca	caa	aga	aat	tat	ttg	gta	caa	6075
Gly	Phe	Leu	Lys	His	Ile	Arg	Ser	Gln	Arg	Asn	Tyr	Leu	Val	Gln	
			1965					1970					1975		
act	gag	gag	caa	tat	gtc	ttc	att	cat	gat	aca	ctg	gtt	gag	gcc	6120
Thr	Glu	Glu	Gln	Tyr	Val	Phe	Ile	His	Asp	Thr	Leu	Val	Glu	Ala	
			1980					1985					1990		
ata	ctt	agt	aaa	gaa	act	gag	gtg	ctg	gac	agt	cat	att	cat	gcc	6165
Ile	Leu	Ser	Lys	Glu	Thr	Glu	Val	Leu	Asp	Ser	His	Ile	His	Ala	
			1995					2000					2005		
tat	gtt	aat	gca	ctc	ctc	att	cct	gga	cca	gca	ggc	aaa	aca	aag	6210
Tyr	Val	Asn	Ala	Leu	Leu	Ile	Pro	Gly	Pro	Ala	Gly	Lys	Thr	Lys	
			2010					2015					2020		
cta	gag	aaa	caa	ttc	cag	ctc	ctg	agc	cag	tca	aat	ata	cag	cag	6255
Leu	Glu	Lys	Gln	Phe	Gln	Leu	Leu	Ser	Gln	Ser	Asn	Ile	Gln	Gln	
			2025					2030					2035		
agt	gac	tat	tct	gca	gcc	cta	aag	caa	tgc	aac	agg	gaa	aag	aat	6300
Ser	Asp	Tyr	Ser	Ala	Ala	Leu	Lys	Gln	Cys	Asn	Arg	Glu	Lys	Asn	
			2040					2045					2050		
cga	act	tct	tct	atc	atc	cct	gtg	gaa	aga	tca	agg	gtt	ggc	att	6345
Arg	Thr	Ser	Ser	Ile	Ile	Pro	Val	Glu	Arg	Ser	Arg	Val	Gly	Ile	
			2055					2060					2065		
tca	tcc	ctg	agt	gga	gaa	ggc	aca	gac	tac	atc	aat	gcc	tcc	tat	6390
Ser	Ser	Leu	Ser	Gly	Glu	Gly	Thr	Asp	Tyr	Ile	Asn	Ala	Ser	Tyr	
			2070					2075					2080		
atc	atg	ggc	tat	tac	cag	agc	aat	gaa	ttc	atc	att	acc	cag	cac	6425
Ile	Met	Gly	Tyr	Tyr	Gln	Ser	Asn	Glu	Phe	Ile	Ile	Thr	Gln	His	
			2085					2090					2095		
cct	ctc	ctt	cat	acc	atc	aag	gat	ttc	tgg	agg	atg	ata	tgg	gac	6480
Pro	Leu	Leu	His	Thr	Ile	Lys	Asp	Phe	Trp	Arg	Met	Ile	Trp	Asp	
			2100					2105					2110		
cat	aat	gcc	caa	ctg	gtg	gtt	atg	att	cct	gat	ggc	caa	aac	atg	6525
His	Asn	Ala	Gln	Leu	Val	Val	Met	Ile	Pro	Asp	Gly	Gln	Asn	Met	
			2115					2120					2125		

atc	atg	ggc	tat	tac	cag	agc	aat	gaa	ttc	atc	att	acc	cag	cac	
Asn	Tyr	Gln	Ser	Phe	Gln	Ser	Asn	Glu	Phe	Ile	Ile	Thr	Gln	His	





```

tataactttt aatacagtag cctgtaaata aaacactctt ccatatgata ttcaacattt 7532
tacaactgca gtattcacct aaagtagaaa taatctgtta cttattgtaa atactgcctt 7592
agtgtctcca tggaccaaat ttatatattat aattgtagat ttttatattt tactactgag 7652
tcaagttttc tagttctgtg taattgttta gtttaatgac gtagttcatt agctgggtctt 7712
actctaccag ttttctgaca ttgtattgtg ttacctaatg cattaacttt gtttcagcat 7772
gtaattttta cttttgtgga aaatagaaat accttcattt tgaaagaagt ttttatgaga 7832
ataacacctt accaaacatt gttcaaatgg tttttatcca aggaattgca aaaataaata 7892
taaattattgc cattaataaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 7941

```

```

<210> 6
<211> 2314
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> gene
<222> (1)..(2314)
<223> PTP-zeta

```

```

<220>
<221> SIGNAL
<222> (1)..(24)
<223> By similarity

```

```

<220>
<221> DOMAIN
<222> (25)..(1635)
<223> Extracellular (potential)

```

```

<220>
<221> MISC_FEATURE
<222> (25)..(2314)
<223> Mature chain; protein-tyrosine phosphatase zeta

```

```

<220>
<221> DOMAIN
<222> (34)..(302)
<223> Carbonic-anhydrase like

```

```

<220>
<221> SITE
<222> (105)..(105)

```

```

<223> Hydroxylation N-linked N-Glycan... potential

```

<220>  
<221> SITE  
<222> (223)..(223)  
<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>  
<221> SITE  
<222> (232)..(232)  
<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>  
<221> SITE  
<222> (324)..(324)  
<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>  
<221> DOMAIN  
<222> (312)..(406)  
<223> Fibronectin Type-III

<220>  
<221> SITE  
<222> (391)..(381)  
<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>  
<221> SITE  
<222> (497)..(497)  
<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>  
<221> SITE  
<222> (501)..(501)  
<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>  
<221> SITE  
<222> (552)..(552)  
<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>  
<221> BINDING  
<222> (587)..(587)  
<223> Chondroitin Sulfate (potential)

<220>  
<221> SITE  
<222> (602)..(602)  
<223> Glycosylation; N-linked (GLCNAC...) (potential)

<221> BINDING  
 <222> (637)..(637)  
 <223> Chondroitin sulfate (potential)  
  
 <220>  
 <221> SITE  
 <222> (677)..(677)  
 <223> Glycosylation; N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> VARSPLIC  
 <222> (755)..(1614)  
 <223> Splicing variant; missing (in short isoform)  
  
 <220>  
 <221> BINDING  
 <222> (997)..(997)  
 <223> Chondroitin sulfate (potential)  
  
 <220>  
 <221> SITE  
 <222> (1017)..(1017)  
 <223> Glycosylation; N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> SITE  
 <222> (1050)..(1050)  
 <223> Glycosylation; N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> SITE  
 <222> (1082)..(1082)  
 <223> Glycosylation; N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> SITE  
 <222> (1122)..(1122)  
 <223> Glycosylation; N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> SITE  
 <222> (1456)..(1456)  
 <223> Glycosylation; N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> BINDING  
 <222> (1548)..(1548)  
 <223> Chondroitin sulfate (potential)  
  
 <220>  
 <221> BINDING

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> SITE

<222> (1617)..(1617)

<223> Glycosylation; N-linked (GLCNAC...) (potential)

<220>

<221> TRANSMEM

<222> (1636)..(1661)

<223> Transmembrane region; potential

<220>

<221> DOMAIN

<222> (1662)..(2314)

<223> Cytoplasmic (potential)

<220>

<221> CONFLICT

<222> (1722)..(1728)

<223> Missing (in ref. 2)

<220>

<221> DOMAIN

<222> (1744)..(1997)

<223> Protein-tyrosine phosphates

<220>

<221> ACT\_SITE

<222> (1932)..(1932)

<223> Active site; by similarity

<220>

<221> DOMAIN

<222> (1998)..(2314)

<223> Protein-tyrosine phosphatase

<220>

<221> MISC\_FEATURE

<222> (2222)..(2222)

<223> Ancestral active site

<100>

<101> Krueger, N.K. and Saito, H.

<102> A human transmembrane protein-tyrosine-phosphatase, PTP zeta, is expressed in brain and has an N-terminal receptor domain homologous to carbonic anhydrases

<103> Proc. Natl. Acad. Sci. USA

<104> 89

<105> 16

<300>  
 <301> Levy, J.B., et al.  
 <302> The cloning of a receptor-type protein tyrosine phosphatase expressed  
 in the central nervous system  
 <303> J. Biol. Chem.  
 <304> 268  
 <305> 14  
 <306> 10573-10581  
 <307> 1993  
 <309>  
 <313> (1)..(2314)

<400> 6

Met Arg Ile Leu Lys Arg Phe Leu Ala Cys Ile Gln Leu Leu Cys Val  
 1 5 10 15

Cys Arg Leu Asp Trp Ala Asn Gly Tyr Tyr Arg Gln Gln Arg Lys Leu  
 20 25 30

Val Glu Glu Ile Gly Trp Ser Tyr Thr Gly Ala Leu Asn Gln Lys Asn  
 35 40 45

Trp Gly Lys Lys Tyr Pro Thr Cys Asn Ser Pro Lys Gln Ser Pro Ile  
 50 55 60

Asn Ile Asp Glu Asp Leu Thr Gln Val Asn Val Asn Leu Lys Lys Leu  
 65 70 75 80

Lys Phe Gln Gly Trp Asp Lys Thr Ser Leu Glu Asn Thr Phe Ile His  
 85 90 95

Asn Thr Gly Lys Thr Val Glu Ile Asn Leu Thr Asn Asp Tyr Arg Val  
 100 105 110

Ser Gly Gly Val Ser Glu Met Val Phe Lys Ala Ser Lys Ile Thr Phe  
 115 120 125

His Trp Gly Lys Cys Asn Met Ser Ser Asp Gly Ser Glu His Ser Leu  
 130 135 140

Arg Ala Leu Ser Ile Leu Phe Glu Val Gly Thr Glu Glu Asn Leu Asp  
180 185 190

Phe Lys Ala Ile Ile Asp Gly Val Glu Ser Val Ser Arg Phe Gly Lys  
195 200 205

Gln Ala Ala Leu Asp Pro Phe Ile Leu Leu Asn Leu Leu Pro Asn Ser  
210 215 220

Thr Asp Lys Tyr Tyr Ile Tyr Asn Gly Ser Leu Thr Ser Pro Pro Cys  
225 230 235 240

Thr Asp Thr Val Asp Trp Ile Val Phe Lys Asp Thr Val Ser Ile Ser  
245 250 255

Glu Ser Gln Leu Ala Val Phe Cys Glu Val Leu Thr Met Gln Gln Ser  
260 265 270

Gly Tyr Val Met Leu Met Asp Tyr Leu Gln Asn Asn Phe Arg Glu Gln  
275 280 285

Gln Tyr Lys Phe Ser Arg Gln Val Phe Ser Ser Tyr Thr Gly Lys Glu  
290 295 300

Glu Ile His Glu Ala Val Cys Ser Ser Glu Pro Glu Asn Val Gln Ala  
305 310 315 320

Asp Pro Glu Asn Tyr Thr Ser Leu Leu Val Thr Trp Glu Arg Pro Arg  
325 330 335

Val Val Tyr Asp Thr Met Ile Glu Lys Phe Ala Val Leu Tyr Gln Gln  
340 345 350

Leu Asp Gly Glu Asp Gln Thr Lys His Glu Phe Leu Thr Asp Gly Tyr  
355 360 365

Gln Asp Leu Gly Ala Ile Leu Asn Asn Leu Leu Pro Asn Met Ser Tyr

Ser Asp Gln Leu Ile Val Asp Met Pro Thr Asp Asn Pro Glu Leu Asp  
405 410 415

Leu Phe Pro Glu Leu Ile Gly Thr Glu Glu Ile Ile Lys Glu Glu Glu  
420 425 430

Glu Gly Lys Asp Ile Glu Glu Gly Ala Ile Val Asn Pro Gly Arg Asp  
435 440 445

Ser Ala Thr Asn Gln Ile Arg Lys Lys Glu Pro Gln Ile Ser Thr Thr  
450 455 460

Thr His Tyr Asn Arg Ile Gly Thr Lys Tyr Asn Glu Ala Lys Thr Asn  
465 470 475 480

Arg Ser Pro Thr Arg Gly Ser Glu Phe Ser Gly Lys Gly Asp Val Pro  
485 490 495

Asn Thr Ser Leu Asn Ser Thr Ser Gln Pro Val Thr Lys Leu Ala Thr  
500 505 510

Glu Lys Asp Ile Ser Leu Thr Ser Gln Thr Val Thr Glu Leu Pro Pro  
515 520 525

His Thr Val Glu Gly Thr Ser Ala Ser Leu Asn Asp Gly Ser Lys Thr  
530 535 540

Val Leu Arg Ser Pro His Met Asn Leu Ser Gly Thr Ala Glu Ser Leu  
545 550 555 560

Asn Thr Val Ser Ile Thr Glu Tyr Glu Glu Glu Ser Leu Leu Thr Ser  
565 570 575

Phe Lys Leu Asp Thr Gly Ala Glu Asp Ser Ser Gly Ser Ser Pro Ala  
580 585 590

Thr Ser Ala Ile Pro Phe Ile Ser Glu Asn Ile Ser Gln Gly Tyr Ile  
595 600 605



Glu Ser Ala Arg Asn Ala Ser Glu Asp Ser Thr Ser Ser Gly Ser Glu  
625 630 635 640

Glu Ser Leu Lys Asp Pro Ser Met Glu Gly Asn Val Trp Phe Pro Ser  
645 650 655

Ser Thr Asp Ile Thr Ala Gln Pro Asp Val Gly Ser Gly Arg Glu Ser  
660 665 670

Phe Leu Gln Thr Asn Tyr Thr Glu Ile Arg Val Asp Glu Ser Glu Lys  
675 680 685

Thr Thr Lys Ser Phe Ser Ala Gly Pro Val Met Ser Gln Gly Pro Ser  
690 695 700

Val Thr Asp Leu Glu Met Pro His Tyr Ser Thr Phe Ala Tyr Phe Pro  
705 710 715 720

Thr Glu Val Thr Pro His Ala Phe Thr Pro Ser Ser Arg Gln Gln Asp  
725 730 735

Leu Val Ser Thr Val Asn Val Val Tyr Ser Gln Thr Thr Gln Pro Val  
740 745 750

Tyr Asn Gly Glu Thr Pro Leu Gln Pro Ser Tyr Ser Ser Glu Val Phe  
755 760 765

Pro Leu Val Thr Pro Leu Leu Leu Asp Asn Gln Ile Leu Asn Thr Thr  
770 775 780

Pro Ala Ala Ser Ser Ser Asp Ser Ala Leu His Ala Thr Pro Val Phe  
785 790 795 800

Pro Ser Val Asp Val Ser Phe Glu Ser Ile Leu Ser Ser Tyr Asp Gly  
805 810 815

Ala Pro Leu Leu Pro Phe Ser Ser Ala Ser Phe Ser Ser Glu Leu Phe  
820 825 830

Thr Ser Ser Arg Tyr Val Thr Ser His Ala Ser Leu Thr Thr Ala Gly

850

855

860

Gly Asp Leu Leu Leu Glu Pro Ser Leu Ala Gln Tyr Ser Asp Val Leu  
865 870 875 880

Ser Thr Thr His Ala Ala Ser Glu Thr Leu Glu Phe Gly Ser Glu Ser  
885 890 895

Gly Val Leu Tyr Lys Thr Leu Met Phe Ser Gln Val Glu Pro Pro Ser  
900 905 910

Ser Asp Ala Met Met His Ala Arg Ser Ser Gly Pro Glu Pro Ser Tyr  
915 920 925

Ala Leu Ser Asp Asn Glu Gly Ser Gln His Ile Phe Thr Val Ser Tyr  
930 935 940

Ser Ser Ala Ile Pro Val His Asp Ser Val Gly Val Thr Tyr Gln Gly  
945 950 955 960

Ser Leu Phe Ser Gly Pro Ser His Ile Pro Ile Pro Lys Ser Ser Leu  
965 970 975

Ile Thr Pro Thr Ala Ser Leu Leu Gln Pro Thr His Ala Leu Ser Gly  
980 985 990

Asp Gly Glu Trp Ser Gly Ala Ser Ser Asp Ser Glu Phe Leu Leu Pro  
995 1000 1005

Asp Thr Asp Gly Leu Thr Ala Leu Asn Ile Ser Ser Pro Val Ser  
1010 1015 1020

Val Ala Glu Phe Thr Tyr Thr Thr Ser Val Phe Gly Asp Asp Asn  
1025 1030 1035

Lys Ala Leu Ser Lys Ser Glu Ile Ile Tyr Gly Asn Glu Thr Glu  
1040 1045 1050

Ser Leu Gln Glu Thr Ser Val Ser Ile Ser Ser Thr Lys Gly Met  
1085 1090 1095

Phe Pro Gly Ser Leu Ala His Thr Thr Thr Lys Val Phe Asp His  
1100 1105 1110

Glu Ile Ser Gln Val Pro Glu Asn Asn Phe Ser Val Gln Pro Thr  
1115 1120 1125

His Thr Val Ser Gln Ala Ser Gly Asp Thr Ser Leu Lys Pro Val  
1130 1135 1140

Leu Ser Ala Asn Ser Glu Pro Ala Ser Ser Asp Pro Ala Ser Ser  
1145 1150 1155

Glu Met Leu Ser Pro Ser Thr Gln Leu Leu Phe Tyr Glu Thr Ser  
1160 1165 1170

Ala Ser Phe Ser Thr Glu Val Leu Leu Gln Pro Ser Phe Gln Ala  
1175 1180 1185

Ser Asp Val Asp Thr Leu Leu Lys Thr Val Leu Pro Ala Val Pro  
1190 1195 1200

Ser Asp Pro Ile Leu Val Glu Thr Pro Lys Val Asp Lys Ile Ser  
1205 1210 1215

Ser Thr Met Leu His Leu Ile Val Ser Asn Ser Ala Ser Ser Glu  
1220 1225 1230

Asn Met Leu His Ser Thr Ser Val Pro Val Phe Asp Val Ser Pro  
1235 1240 1245

Thr Ser His Met His Ser Ala Ser Leu Gln Gly Leu Thr Ile Ser  
1250 1255 1260

Tyr Ala Ser Glu Lys Tyr Glu Pro Val Leu Leu Lys Ser Glu Ser

Gln Thr Ala Asn Leu Glu Ile Asn Gln Ala His Pro Pro Lys Gly  
1295 1300 1305

Arg His Val Phe Ala Thr Pro Val Leu Ser Ile Asp Glu Pro Leu  
1310 1315 1320

Asn Thr Leu Ile Asn Lys Leu Ile His Ser Asp Glu Ile Leu Thr  
1325 1330 1335

Ser Thr Lys Ser Ser Val Thr Gly Lys Val Phe Ala Gly Ile Pro  
1340 1345 1350

Thr Val Ala Ser Asp Thr Phe Val Ser Thr Asp His Ser Val Pro  
1355 1360 1365

Ile Gly Asn Gly His Val Ala Ile Thr Ala Val Ser Pro His Arg  
1370 1375 1380

Asp Gly Ser Val Thr Ser Thr Lys Leu Leu Phe Pro Ser Lys Ala  
1385 1390 1395

Thr Ser Glu Leu Ser His Ser Ala Lys Ser Asp Ala Gly Leu Val  
1400 1405 1410

Gly Gly Gly Glu Asp Gly Asp Thr Asp Asp Asp Gly Asp Asp Asp  
1415 1420 1425

Asp Asp Arg Asp Ser Asp Gly Leu Ser Ile His Lys Cys Met Ser  
1430 1435 1440

Cys Ser Ser Tyr Arg Glu Ser Gln Glu Lys Val Met Asn Asp Ser  
1445 1450 1455

Asp Thr His Glu Asn Ser Leu Met Asp Gln Asn Asn Pro Ile Ser  
1460 1465 1470

Tyr Ser Leu Ser Glu Asn Ser Glu Glu Asp Asn Arg Val Thr Ser  
1475 1480 1485



1715		1720		1725
Glu Val Gln Ser Cys Thr Val Asp Leu Gly Ile Thr Ala Asp Ser				
1730		1735		1740
Ser Asn His Pro Asp Asn Lys His Lys Asn Arg Tyr Ile Asn Ile				
1745		1750		1755
Val Ala Tyr Asp His Ser Arg Val Lys Leu Ala Gln Leu Ala Glu				
1760		1765		1770
Lys Asp Gly Lys Leu Thr Asp Tyr Ile Asn Ala Asn Tyr Val Asp				
1775		1780		1785
Gly Tyr Asn Arg Pro Lys Ala Tyr Ile Ala Ala Gln Gly Pro Leu				
1790		1795		1800
Lys Ser Thr Ala Glu Asp Phe Trp Arg Met Ile Trp Glu His Asn				
1805		1810		1815
Val Glu Val Ile Val Met Ile Thr Asn Leu Val Glu Lys Gly Arg				
1820		1825		1830
Arg Lys Cys Asp Gln Tyr Trp Pro Ala Asp Gly Ser Glu Glu Tyr				
1835		1840		1845
Gly Asn Phe Leu Val Thr Gln Lys Ser Val Gln Val Leu Ala Tyr				
1850		1855		1860
Tyr Thr Val Arg Asn Phe Thr Leu Arg Asn Thr Lys Ile Lys Lys				
1865		1870		1875
Gly Ser Gln Lys Gly Arg Pro Ser Gly Arg Val Val Thr Gln Tyr				
1880		1885		1890
His Tyr Thr Gln Trp Pro Asp Met Gly Val Pro Glu Tyr Ser Leu				
1895		1900		1905

1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995

Thr Gly Thr Tyr Ile Val Leu Asp Ser Met Leu Gln Gln Ile Gln  
1940 1945 1950

His Glu Gly Thr Val Asn Ile Phe Gly Phe Leu Lys His Ile Arg  
1955 1960 1965

Ser Gln Arg Asn Tyr Leu Val Gln Thr Glu Glu Gln Tyr Val Phe  
1970 1975 1980

Ile His Asp Thr Leu Val Glu Ala Ile Leu Ser Lys Glu Thr Glu  
1985 1990 1995

Val Leu Asp Ser His Ile His Ala Tyr Val Asn Ala Leu Leu Ile  
2000 2005 2010

Pro Gly Pro Ala Gly Lys Thr Lys Leu Glu Lys Gln Phe Gln Leu  
2015 2020 2025

Leu Ser Gln Ser Asn Ile Gln Gln Ser Asp Tyr Ser Ala Ala Leu  
2030 2035 2040

Lys Gln Cys Asn Arg Glu Lys Asn Arg Thr Ser Ser Ile Ile Pro  
2045 2050 2055

Val Glu Arg Ser Arg Val Gly Ile Ser Ser Leu Ser Gly Glu Gly  
2060 2065 2070

Thr Asp Tyr Ile Asn Ala Ser Tyr Ile Met Gly Tyr Tyr Gln Ser  
2075 2080 2085

Asn Glu Phe Ile Ile Thr Gln His Pro Leu Leu His Thr Ile Lys  
2090 2095 2100

Asp Phe Trp Arg Met Ile Trp Asp His Asn Ala Gln Leu Val Val  
2105 2110 2115

Met Ile Pro Asp Gly Gln Asn Met Ala Glu Asp Glu Phe Val Tyr

Thr Leu Met Ala Glu Glu His Lys Cys Leu Ser Asn Glu Glu Lys  
2150 2155 2160

Leu Ile Ile Gln Asp Phe Ile Leu Glu Ala Thr Gln Asp Asp Tyr  
2165 2170 2175

Val Leu Glu Val Arg His Phe Gln Cys Pro Lys Trp Pro Asn Pro  
2180 2185 2190

Asp Ser Pro Ile Ser Lys Thr Phe Glu Leu Ile Ser Val Ile Lys  
2195 2200 2205

Glu Glu Ala Ala Asn Arg Asp Gly Pro Met Ile Val His Asp Glu  
2210 2215 2220

His Gly Gly Val Thr Ala Gly Thr Phe Cys Ala Leu Thr Thr Leu  
2225 2230 2235

Met His Gln Leu Glu Lys Glu Asn Ser Val Asp Val Tyr Gln Val  
2240 2245 2250

Ala Lys Met Ile Asn Leu Met Arg Pro Gly Val Phe Ala Asp Ile  
2255 2260 2265

Glu Gln Tyr Gln Phe Leu Tyr Lys Val Ile Leu Ser Leu Val Ser  
2270 2275 2280

Thr Arg Gln Glu Glu Asn Pro Ser Thr Ser Leu Asp Ser Asn Gly  
2285 2290 2295

Ala Ala Leu Pro Asp Gly Asn Ile Ala Glu Ser Leu Glu Ser Leu  
2300 2305 2310

Val

<210> 7  
<211> 1518

...  
... 1 1 1 1 1



<223> Angiopoietin-like 2 (ANGPTL2), mRNA

<220>

<221> CDS

<222> (22)..(1503)

<223>

<300>

<301> Kim, I., Moon, S.O., Koh, K.N., Kim, H., Uhm, C.S., Kwak, H.J., Kim, N.G. and Koh, G.Y.

<302> Molecular cloning, expression, and characterization of angiopoietin-related protein. angiopoietin-related protein induces endothelial cell sprouting

<303> J. Biol. Chem.

<304> 274

<305> 37

<306> 26523-26528

<307> 1999

<309>

<313> (1)..(1518)

<400> 7

aaccaccatt ttgcaaggac c atg agg cca ctg tgc gtg aca tgc tgg tgg	51
Met Arg Pro Leu Cys Val Thr Cys Trp Trp	
1 5 10	

ctc gga ctg ctg gct gcc atg gga gct gtt gca ggc cag gag gac ggt	99
Leu Gly Leu Leu Ala Ala Met Gly Ala Val Ala Gly Gln Glu Asp Gly	
15 20 25	

ttt gag ggc act gag gag ggc tcg cca aga gag ttc att tac cta aac	147
Phe Glu Gly Thr Glu Glu Gly Ser Pro Arg Glu Phe Ile Tyr Leu Asn	
30 35 40	

agg tac aag cgg gcg ggc gag tcc cag gac aag tgc acc tac acc ttc	195
Arg Tyr Lys Arg Ala Gly Glu Ser Gln Asp Lys Cys Thr Tyr Thr Phe	
45 50 55	

att gtg ccc cag cag cgg gtc acg ggt gcc atc tgc gtc aac tcc aag	243
Ile Val Pro Gln Gln Arg Val Thr Gly Ala Ile Cys Val Asn Ser Lys	
60 65 70	

gag cct gag gtg ctt ctg gag aac cga gtg cat aag cag gag cta gag	291
Glu Pro Glu Val Leu Leu Glu Asn Arg Val His Lys Gln Glu Leu Glu	
75 80 85 90	

ctg ctc aac aat gag ctg ctc aag cag aag cgg cag atc gag aca ctg	339
Leu Leu Asn Asn Glu Leu Leu Lys Gln Lys Arg Gln Ile Glu Thr Leu	
95 100 105	

Leu Arg Lys His Ser Arg Asn Met Asn Ser Arg Val Thr Gln Leu Tyr

125	130	135	
atg cag ctc ctg cac gag atc atc cgc aag cgg gac aac gcg ttg gag Met Gln Leu Leu His Glu Ile Ile Arg Lys Arg Asp Asn Ala Leu Glu 140 145 150			483
ctc tcc cag ctg gag aac agg atc ctg aac cag aca gcc gac atg ctg Leu Ser Gln Leu Glu Asn Arg Ile Leu Asn Gln Thr Ala Asp Met Leu 155 160 165 170			531
cag ctg gcc agc aag tac aag gac ctg gag cac aag tac cag cac ctg Gln Leu Ala Ser Lys Tyr Lys Asp Leu Glu His Lys Tyr Gln His Leu 175 180 185			579
gcc aca ctg gcc cac aac caa tca gag atc atc gcg cag ctt gag gag Ala Thr Leu Ala His Asn Gln Ser Glu Ile Ile Ala Gln Leu Glu Glu 190 195 200			627
cac tgc cag agg gtg ccc tgc gcc agg ccc gtc ccc cag cca ccc ccc His Cys Gln Arg Val Pro Ser Ala Arg Pro Val Pro Gln Pro Pro Pro 205 210 215			675
gct gcc ccg ccc cgg gtc tac caa cca ccc acc tac aac cgc atc atc Ala Ala Pro Pro Arg Val Tyr Gln Pro Pro Thr Tyr Asn Arg Ile Ile 220 225 230			723
aac cag atc tct acc aac gag atc cag agt gac cag aac ctg aag gtg Asn Gln Ile Ser Thr Asn Glu Ile Gln Ser Asp Gln Asn Leu Lys Val 235 240 245 250			771
ctg cca ccc cct ctg ccc act atg ccc act ctc acc agc ctc cca tct Leu Pro Pro Pro Leu Pro Thr Met Pro Thr Leu Thr Ser Leu Pro Ser 255 260 265			819
tcc acc gac aag ccg tgc gcc cca tgg aga gac tgc ctg cag gcc ctg Ser Thr Asp Lys Pro Ser Gly Pro Trp Arg Asp Cys Leu Gln Ala Leu 270 275 280			867
gag gat gcc cac gac acc agc tcc atc tac ctg gtg aag ccg gag aac Glu Asp Gly His Asp Thr Ser Ser Ile Tyr Leu Val Lys Pro Glu Asn 285 290 295			915
acc aac cgc ctc atg cag gtg tgg tgc gac cag aga cac gac ccc ggg Thr Asn Arg Leu Met Gln Val Trp Cys Asp Gln Arg His Asp Pro Gly 300 305 310			963
ggc tgg acc gtc atc cag aga cgc ctg gat gcc tct gtt aac ttc ttc Gly Trp Thr Val Ile Gln Arg Arg Leu Asp Gly Ser Val Asn Phe Phe 315 320 325 330			1011
agg aac tgg aag acc tac aag caa gga ttt ggg aac att gat ggc gaa			1059

tac aaa ctc ctg gtg acc atg gag gac tgg tcc ggc cgc aaa gtc ttt	1155
Tyr Lys Leu Leu Val Thr Met Glu Asp Trp Ser Gly Arg Lys Val Phe	
365 370 375	
gca gaa tac gcc agt ttc cgc ctg gaa cct gag agc gag tat tat aag	1203
Ala Glu Tyr Ala Ser Phe Arg Leu Glu Pro Glu Ser Glu Tyr Tyr Lys	
380 385 390	
ctg cgg ctg ggg cgc tac cat ggc aat gcg ggt gac tcc ttt aca tgg	1251
Leu Arg Leu Gly Arg Tyr His Gly Asn Ala Gly Asp Ser Phe Thr Trp	
395 400 405 410	
cac aac ggc aag cag ttc acc acc ctg gac aga gat cat gat gtc tac	1299
His Asn Gly Lys Gln Phe Thr Thr Leu Asp Arg Asp His Asp Val Tyr	
415 420 425	
aca gga aac tgt gcc cac tac cag aag gga ggc tgg tgg tat aac gcc	1347
Thr Gly Asn Cys Ala His Tyr Gln Lys Gly Gly Trp Trp Tyr Asn Ala	
430 435 440	
tgt gcc cac tcc aac ctc aac ggg gtc tgg tac cgc ggg ggc cat tac	1395
Cys Ala His Ser Asn Leu Asn Gly Val Trp Tyr Arg Gly Gly His Tyr	
445 450 455	
cgg agc cgc tac cag gac gga gtc tac tgg gct gag ttc cga gga ggc	1443
Arg Ser Arg Tyr Gln Asp Gly Val Tyr Trp Ala Glu Phe Arg Gly Gly	
460 465 470	
tct tac tca ctc aag aaa gtg gtg atg atg atc cga ccg aac ccc aac	1491
Ser Tyr Ser Leu Lys Lys Val Val Met Met Ile Arg Pro Asn Pro Asn	
475 480 485 490	
acc ttc cac taa gccagctccc cctcc	1518
Thr Phe His	

<210> 8  
 <211> 493  
 <212> PRT  
 <213> Homo sapiens  
  
 <220>  
 <221> gene  
 <222> (1)..(493)  
 <223> Angiopoietin-like 2 (ANGPTL2), protein  
  
 <220>  
 <221> SIGNAL  
 <222> (1)..(22)

...  
 ... Angiopoietin-related protein ...

<220>  
 <221> DOMAIN  
 <222> (76)..(115)  
 <223> Coiled Coil (potential)

<220>  
 <221> DOMAIN  
 <222> (152)..(206)  
 <223> Coiled Coil (potential)

<220>  
 <221> DOMAIN  
 <222> (438)..(450)  
 <223> Fibrinogen C-terminal

<220>  
 <221> CARBOHYD  
 <222> (164)..(164)  
 <223> N-linked (GLCNAC...) (potential)

<220>  
 <221> CARBOHYD  
 <222> (192)..(192)  
 <223> N-linked (GLCNAC...) (potential)

<400> 8

Met	Arg	Pro	Leu	Cys	Val	Thr	Cys	Trp	Trp	Leu	Gly	Leu	Leu	Ala	Ala
1				5				10						15	

Met	Gly	Ala	Val	Ala	Gly	Gln	Glu	Asp	Gly	Phe	Glu	Gly	Thr	Glu	Glu
			20					25					30		

Gly	Ser	Pro	Arg	Glu	Phe	Ile	Tyr	Leu	Asn	Arg	Tyr	Lys	Arg	Ala	Gly
			35				40					45			

Glu	Ser	Gln	Asp	Lys	Cys	Thr	Tyr	Thr	Phe	Ile	Val	Pro	Gln	Gln	Arg
		50				55					60				

Val	Thr	Gly	Ala	Ile	Cys	Val	Asn	Ser	Lys	Glu	Pro	Glu	Val	Leu	Leu
65					70					75				80	

Glu	Asn	Arg	Val	His	Lys	Gln	Glu	Leu	Glu	Leu	Leu	Asn	Asn	Glu	Leu
				85				90						95	

Asp Gly Gly Ile Val Ser Glu Val Lys Leu Leu Arg Lys Glu Ser Arg  
115 120 125

Asn Met Asn Ser Arg Val Thr Gln Leu Tyr Met Gln Leu Leu His Glu  
130 135 140

Ile Ile Arg Lys Arg Asp Asn Ala Leu Glu Leu Ser Gln Leu Glu Asn  
145 150 155 160

Arg Ile Leu Asn Gln Thr Ala Asp Met Leu Gln Leu Ala Ser Lys Tyr  
165 170 175

Lys Asp Leu Glu His Lys Tyr Gln His Leu Ala Thr Leu Ala His Asn  
180 185 190

Gln Ser Glu Ile Ile Ala Gln Leu Glu Glu His Cys Gln Arg Val Pro  
195 200 205

Ser Ala Arg Pro Val Pro Gln Pro Pro Pro Ala Ala Pro Pro Arg Val  
210 215 220

Tyr Gln Pro Pro Thr Tyr Asn Arg Ile Ile Asn Gln Ile Ser Thr Asn  
225 230 235 240

Glu Ile Gln Ser Asp Gln Asn Leu Lys Val Leu Pro Pro Pro Leu Pro  
245 250 255

Thr Met Pro Thr Leu Thr Ser Leu Pro Ser Ser Thr Asp Lys Pro Ser  
260 265 270

Gly Pro Trp Arg Asp Cys Leu Gln Ala Leu Glu Asp Gly His Asp Thr  
275 280 285

Ser Ser Ile Tyr Leu Val Lys Pro Glu Asn Thr Asn Arg Leu Met Gln  
290 295 300

Val Trp Cys Asp Gln Arg His Asp Pro Gly Gly Trp Thr Val Ile Gln  
305 310 315 320

Lys Asn Gly Ile Gly Asn Ile Arg Gly Glu Tyr Trp Leu Gly Leu Asn

340

345

350

Asn Ile Tyr Trp Leu Thr Asn Gln Gly Asn Tyr Lys Leu Leu Val Thr  
 355 360 365

Met Glu Asp Trp Ser Gly Arg Lys Val Phe Ala Glu Tyr Ala Ser Phe  
 370 375 380

Arg Leu Glu Pro Glu Ser Glu Tyr Tyr Lys Leu Arg Leu Gly Arg Tyr  
 385 390 395 400

His Gly Asn Ala Gly Asp Ser Phe Thr Trp His Asn Gly Lys Gln Phe  
 405 410 415

Thr Thr Leu Asp Arg Asp His Asp Val Tyr Thr Gly Asn Cys Ala His  
 420 425 430

Tyr Gln Lys Gly Gly Trp Trp Tyr Asn Ala Cys Ala His Ser Asn Leu  
 435 440 445

Asn Gly Val Trp Tyr Arg Gly Gly His Tyr Arg Ser Arg Tyr Gln Asp  
 450 455 460

Gly Val Tyr Trp Ala Glu Phe Arg Gly Gly Ser Tyr Ser Leu Lys Lys  
 465 470 475 480

Val Val Met Met Ile Arg Pro Asn Pro Asn Thr Phe His  
 485 490

<210> 9  
 <211> 2133  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> gene  
 <222> (1)..(2133)  
 <223> Human SPARC/osteonectin mRNA

<220>  
 <221> CDS

atg agg gcc tgg atc ttc ttt ctc ctt tgc ctg gcc ggg agg gcc ttg	105
Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu	
1 5 10 15	
gca gcc cct cag caa gaa gcc ctg cct gat gag aca gag gtg gtg gaa	153
Ala Ala Pro Gln Gln Glu Ala Leu Pro Asp Glu Thr Glu Val Val Glu	
20 25 30	
gaa act gtg gca gag gtg act gag gta tct gtg gga gct aat cct gtc	201
Glu Thr Val Ala Glu Val Thr Glu Val Ser Val Gly Ala Asn Pro Val	
35 40 45	
cag gtg gaa gta gga gaa ttt gat gat ggt gca gag gaa acc gaa gag	249
Gln Val Glu Val Gly Glu Phe Asp Asp Gly Ala Glu Glu Thr Glu Glu	
50 55 60	
gag gtg gtg gcg gaa aat ccc tgc cag aac cac cac tgc aaa cac ggc	297
Glu Val Val Ala Glu Asn Pro Cys Gln Asn His His Cys Lys His Gly	
65 70 75 80	
aag gtg tgc gag ctg gat gag aac aac acc ccc atg tgc gtg tgc cag	345
Lys Val Cys Glu Leu Asp Glu Asn Asn Thr Pro Met Cys Val Cys Gln	
85 90 95	
gac ccc acc agc tgc cca gcc ccc att ggc gag ttt gag aag gtg tgc	393
Asp Pro Thr Ser Cys Pro Ala Pro Ile Gly Glu Phe Glu Lys Val Cys	
100 105 110	
agc aat gac aac aag acc ttc gac tct tcc tgc cac ttc ttt gcc aca	441
Ser Asn Asp Asn Lys Thr Phe Asp Ser Ser Cys His Phe Phe Ala Thr	
115 120 125	
aag tgc acc ctg gag ggc acc aag aag ggc cac aag ctc cac ctg gac	489
Lys Cys Thr Leu Glu Gly Thr Lys Lys Gly His Lys Leu His Leu Asp	
130 135 140	
tac atc ggg cct tgc aaa tac atc ccc cct tgc ctg gac tct gag ctg	537
Tyr Ile Gly Pro Cys Lys Tyr Ile Pro Pro Cys Leu Asp Ser Glu Leu	
145 150 155 160	
acc gaa ttc ccc ctg cgc atg cgg gac tgg ctc aag aac gtc ctg gtc	585
Thr Glu Phe Pro Leu Arg Met Arg Asp Trp Leu Lys Asn Val Leu Val	
165 170 175	
acc ctg tat gag agg gat gag gac aac aac ctt ctg act gag aag cag	633
Thr Leu Tyr Glu Arg Asp Glu Asp Asn Asn Leu Leu Thr Glu Lys Gln	
180 185 190	
aag ctg cgg gtg aag aag atc cat gag aat gag aag cgc ctg gag gca	681
Lys Leu Arg Val Lys Lys Ile His Glu Asn Glu Lys Arg Leu Glu Ala	
195 200 205	





```

<210> 10
<211> 303
<212> PRT
<213> Homo sapiens

<220>
<221> SIGNAL
<222> (1)..(17)
<223>

<220>
<221> gene
<222> (1)..(303)
<223> SPARC, osteonectin

<220>
<221> CHAIN
<222> (18)..(303)
<223> SPARC

<220>
<221> DOMAIN
<222> (22)..(69)
<223> ASP/GLU-RICH (ACIDIC, BINDS CALCIUM)

<220>
<221> DOMAIN
<222> (94)..(149)
<223> KAZAL-LIKE

<220>
<221> DOMAIN
<222> (185)..(221)
<223> ALPHA-HELIX CONFORMATION (PROBABLE)

<220>
<221> CA_BIND
<222> (274)..(285)
<223> 12 EF-HAND

<220>
<221> CARBOHYD
<222> (116)..(116)
<223> N-LINKED (GLCNAC ...) (PROBABLE)

<220>
<221> DISULFID
<222> (72)..(83)

```

<220>  
<221> DISULFID  
<222> (95)..(130)  
<223>

<220>  
<221> DISULFID  
<222> (101)..(123)  
<223>

<220>  
<221> DISULFID  
<222> (112)..(149)  
<223>

<220>  
<221> DISULFID  
<222> (155)..(265)  
<223>

<220>  
<221> DISULFID  
<222> (273)..(289)  
<223>

<400> 10

Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu  
1 5 10 15

Ala Ala Pro Gln Gln Glu Ala Leu Pro Asp Glu Thr Glu Val Val Glu  
20 25 30

Glu Thr Val Ala Glu Val Thr Glu Val Ser Val Gly Ala Asn Pro Val  
35 40 45

Gln Val Glu Val Gly Glu Phe Asp Asp Gly Ala Glu Glu Thr Glu Glu  
50 55 60

Glu Val Val Ala Glu Asn Pro Cys Gln Asn His His Cys Lys His Gly  
65 70 75 80

Lys Val Cys Glu Leu Asp Glu Asn Asn Thr Pro Met Cys Val Cys Gln  
85 90 95

Ser Asn Asp Asn Lys Thr Phe Asp Ser Ser Cys His Phe Phe Ala Thr  
115 120 125

Lys Cys Thr Leu Glu Gly Thr Lys Lys Gly His Lys Leu His Leu Asp  
130 135 140

Tyr Ile Gly Pro Cys Lys Tyr Ile Pro Pro Cys Leu Asp Ser Glu Leu  
145 150 155 160

Thr Glu Phe Pro Leu Arg Met Arg Asp Trp Leu Lys Asn Val Leu Val  
165 170 175

Thr Leu Tyr Glu Arg Asp Glu Asp Asn Asn Leu Leu Thr Glu Lys Gln  
180 185 190

Lys Leu Arg Val Lys Lys Ile His Glu Asn Glu Lys Arg Leu Glu Ala  
195 200 205

Gly Asp His Pro Val Glu Leu Leu Ala Arg Asp Phe Glu Lys Asn Tyr  
210 215 220

Asn Met Tyr Ile Phe Pro Val His Trp Gln Phe Gly Gln Leu Asp Gln  
225 230 235 240

His Pro Ile Asp Gly Tyr Leu Ser His Thr Glu Leu Ala Pro Leu Arg  
245 250 255

Ala Pro Leu Ile Pro Met Glu His Cys Thr Thr Arg Phe Phe Glu Thr  
260 265 270

Cys Asp Leu Asp Asn Asp Lys Tyr Ile Ala Leu Asp Glu Trp Ala Gly  
275 280 285

Cys Phe Gly Ile Lys Gln Lys Asp Ile Asp Lys Asp Leu Val Ile  
290 295 300

<210> 11  
<211> 4586  
<212> DNA

...  
... DNA ...

<220>  
 <221> CDS  
 <222> (195)..(4367)  
 <223>

<300>  
 <301> Park,M., Dean,M., Kaul,K., Braun,M.J., Gonda,M.A. and Vande Woude,G.  
 <302> Sequence of MET protooncogene cDNA has features characteristic of the  
 tyrosine kinase family of growth-factor receptors  
 <303> Proc. Natl. Acad. Sci. U.S.A.  
 <304> 84  
 <305> 18  
 <306> 6379-6383  
 <307> 1987  
 <309>  
 <313> (1)..(4586)

<400> 11  
 gaattccgcc ctgcgcgcgc ggggcgcgc gagcgctttg tgagcagatg cggagccgag 60  
 tggaggggcgc gagccagatg cggggcgaca gctgacttgc tgagaggagg cggggaggcg 120  
 cggagcgcgc gtgtggctct tgcgcgcgtg acttctccac tggttcctgg gcaccgaaag 180  
 ataaacctct cata atg aag gcc ccc gct gtg ctt gca cct ggc atc ctc 230  
 Met Lys Ala Pro Ala Val Leu Ala Pro Gly Ile Leu  
 1 5 10  
 gtg ctc ctg ttt acc ttg gtg cag agg agc aat ggg gag tgt aaa gag 278  
 Val Leu Leu Phe Thr Leu Val Gln Arg Ser Asn Gly Glu Cys Lys Glu  
 15 20 25  
 gca cta gca aag tcc gag atg aat gtg aat atg aag tat cag ctt ccc 326  
 Ala Leu Ala Lys Ser Glu Met Asn Val Asn Met Lys Tyr Gln Leu Pro  
 30 35 40  
 aac ttc acc ggc gaa aca ccc atc cag aat gtc att cta cat gag cat 374  
 Asn Phe Thr Ala Glu Thr Pro Ile Gln Asn Val Ile Leu His Glu His  
 45 50 55 60  
 cac att ttc ctt ggt gcc act aac tac att tat gtt tta aat gag gaa 422  
 His Ile Phe Leu Gly Ala Thr Asn Tyr Ile Tyr Val Leu Asn Glu Glu  
 65 70 75  
 gac ctt cag aag gtt gct gag tac aag act ggg cct gtg ctg gaa cac 470  
 Asp Leu Gln Lys Val Ala Glu Tyr Lys Thr Gly Pro Val Leu Glu His  
 80 85 90

111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

Gly	Gly	Val	Trp	Lys	Asp	Asn	Ile	Asn	Met	Ala	Leu	Val	Val	Asp	Thr		
110						115					120						
tac	tat	gat	gat	aaa	ctc	att	agc	tgt	ggc	agc	gtc	aac	aga	ggg	acc	614	
Tyr	Tyr	Asp	Asp	Gln	Leu	Ile	Ser	Cys	Gly	Ser	Val	Asn	Arg	Gly	Thr		
125					130					135					140		
tgc	cag	cga	cat	gtc	ttt	ccc	cac	aat	cat	act	gct	gac	ata	cag	tcg	662	
Cys	Gln	Arg	His	Val	Phe	Pro	His	Asn	His	Thr	Ala	Asp	Ile	Gln	Ser		
				145					150						155		
gag	gtt	cac	tgc	ata	ttc	tcc	cca	cag	ata	gaa	gag	ccc	agc	cag	tgt	710	
Glu	Val	His	Cys	Ile	Phe	Ser	Pro	Gln	Ile	Glu	Glu	Pro	Ser	Gln	Cys		
			160					165					170				
cct	gac	tgt	gtg	gtg	agc	gcc	ctg	gga	gcc	aaa	gtc	ctt	tca	tct	gta	758	
Pro	Asp	Cys	Val	Val	Ser	Ala	Leu	Gly	Ala	Lys	Val	Leu	Ser	Ser	Val		
		175					180						185				
aag	gac	cgg	ttc	atc	aac	ttc	ttt	gta	ggc	aat	acc	ata	aat	tct	tct	806	
Lys	Asp	Arg	Phe	Ile	Asn	Phe	Phe	Val	Gly	Asn	Thr	Ile	Asn	Ser	Ser		
	190					195					200						
tat	ttc	cca	gat	cat	cca	ttg	cat	tcg	ata	tca	gtg	aga	agg	cta	aag	854	
Tyr	Phe	Pro	Asp	His	Pro	Leu	His	Ser	Ile	Ser	Val	Arg	Arg	Leu	Lys		
205					210				215						220		
gaa	acg	aaa	gat	ggg	ttt	atg	ttt	ttg	acg	gac	cag	tcc	tac	att	gat	902	
Glu	Thr	Lys	Asp	Gly	Phe	Met	Phe	Leu	Thr	Asp	Gln	Ser	Tyr	Ile	Asp		
			225					230						235			
gtt	tta	cct	gag	ttc	aga	gat	tct	tac	ccc	att	aag	tat	gtc	cat	gcc	950	
Val	Leu	Pro	Glu	Phe	Arg	Asp	Ser	Tyr	Pro	Ile	Lys	Tyr	Val	His	Ala		
		240						245					250				
ttt	gaa	agc	aac	aat	ttt	att	tac	ttc	ttg	acg	gtc	caa	agg	gaa	act	998	
Phe	Glu	Ser	Asn	Asn	Phe	Ile	Tyr	Phe	Leu	Thr	Val	Gln	Arg	Glu	Thr		
		255					260					265					
cta	gat	gct	cag	act	ttt	cac	aca	aga	ata	atc	agg	ttc	tgt	tcc	ata	1046	
Leu	Asp	Ala	Gln	Thr	Phe	His	Thr	Arg	Ile	Ile	Arg	Phe	Cys	Ser	Ile		
	270					275					280						
aac	tct	gga	ttg	cat	tcc	tac	atg	gaa	atg	cct	ctg	gag	tgt	att	ctc	1094	
Asn	Ser	Gly	Leu	His	Ser	Tyr	Met	Glu	Met	Pro	Leu	Glu	Cys	Ile	Leu		
285				290						295					300		
aca	gaa	aag	aga	aaa	aag	aga	tcc	aca	aag	aag	gaa	gtg	ttt	aat	ata	1142	
Thr	Glu	Lys	Arg	Lys	Lys	Arg	Ser	Thr	Lys	Lys	Glu	Val	Phe	Asn	Ile		
			305					310						315			

115 Gly Ala Val Leu Asn Arg Arg 120 Leu Phe Gly Val Ile Ala Thr

335	340	345	
agc aag cca gat tct gcc gaa cca atg gat cga tct gcc atg tgt gca Ser Lys Pro Asp Ser Ala Glu Pro Met Asp Arg Ser Ala Met Cys Ala 350 355 360			1286
ttc cct atc aaa tat gtc aac gac ttc ttc aac aag atc gtc aac aaa Phe Pro Ile Lys Tyr Val Asn Asp Phe Phe Asn Lys Ile Val Asn Lys 365 370 375 380			1334
aac aat gtg aga tgt ctc cag cat ttt tac gga ccc aat cat gag cac Asn Asn Val Arg Cys Leu Gln His Phe Tyr Gly Pro Asn His Glu His 385 390 395			1382
tgc ttt aat agg aca ctt ctg aga aat tca tca ggc tgt gaa gcg cgc Cys Phe Asn Arg Thr Leu Leu Arg Asn Ser Ser Gly Cys Glu Ala Arg 400 405 410			1430
cgt gat gaa tat cga aca gag ttt acc aca gct ttg cag cgc gtt gac Arg Asp Glu Tyr Arg Thr Glu Phe Thr Thr Ala Leu Gln Arg Val Asp 415 420 425			1478
tta ttc atg ggt caa ttc agc gaa gtc ctc tta aca tct ata tcc acc Leu Phe Met Gly Gln Phe Ser Glu Val Leu Leu Thr Ser Ile Ser Thr 430 435 440			1526
ttc att aaa gga gac ctc acc ata gct aat ctt ggg aca tca gag ggt Phe Ile Lys Gly Asp Leu Thr Ile Ala Asn Leu Gly Thr Ser Glu Gly 445 450 455 460			1574
cgc ttc atg cag gtt gtg gtt tct cga tca gga cca tca acc cct cat Arg Phe Met Gln Val Val Ser Arg Ser Gly Pro Ser Thr Pro His 465 470 475			1622
gtg aat ttt ctc ctg gac tcc cat cca gtg tct cca gaa gtg att gtg Val Asn Phe Leu Leu Asp Ser His Pro Val Ser Pro Glu Val Ile Val 480 485 490			1670
gag cat aca tta aac caa aat ggc tac aca ctg gtt atc act ggg aag Glu His Thr Leu Asn Gln Asn Gly Tyr Thr Leu Val Ile Thr Gly Lys 495 500 505			1718
aag atc acg aag atc cca ttg aat ggc ttg ggc tgc aga cat ttc cag Lys Ile Thr Lys Ile Pro Leu Asn Gly Leu Gly Cys Arg His Phe Gln 510 515 520			1766
tcg tgc agt caa tgc ctc tct gcc cca ccc ttt gtt cag tgt ggc tgg Ser Cys Ser Gln Cys Leu Ser Ala Pro Pro Phe Val Gln Cys Gly Trp 525 530 535 540			1814
tgc cac gac aaa tgt gtg cga tgc gag gaa tgc ctg agc ggg aca tgg			1862

gca ccc ctt gaa gga ggg aca agg ctg acc ata tgt ggc tgg gac ttt Ala Pro Leu Glu Gly Gly Thr Arg Leu Thr Ile Cys Gly Trp Asp Phe 575 580 585	1958
gga ttt cgg agg aat aat aaa ttt gat tta aag aaa act aga gtt ctc Gly Phe Arg Arg Asn Asn Lys Phe Asp Leu Lys Lys Thr Arg Val Leu 590 595 600	2006
ctt gga aat gag agc tgc acc ttg act tta agt gag agc acg atg aat Leu Gly Asn Glu Ser Cys Thr Leu Thr Leu Ser Glu Ser Thr Met Asn 605 610 615 620	2054
aca ttg aaa tgc aca gtt ggt cct gcc atg aat aag cat ttc aat atg Thr Leu Lys Cys Thr Val Gly Pro Ala Met Asn Lys His Phe Asn Met 625 630 635	2102
tcc ata att att tca aat ggc cac ggg aca aca caa tac agt aca ttc Ser Ile Ile Ile Ser Asn Gly His Gly Thr Thr Gln Tyr Ser Thr Phe 640 645 650	2150
tcc tat gtg gat cct gta ata aca agt att tcg ccg aaa tac ggt cct Ser Tyr Val Asp Pro Val Ile Thr Ser Ile Ser Pro Lys Tyr Gly Pro 655 660 665	2198
atg gct ggt ggc act tta ctt act tta act gga aat tac cta aac agt Met Ala Gly Gly Thr Leu Leu Thr Leu Thr Gly Asn Tyr Leu Asn Ser 670 675 680	2246
ggg aat tct aga cac att tca att ggt gga aaa aca tgt act tta aaa Gly Asn Ser Arg His Ile Ser Ile Gly Gly Lys Thr Cys Thr Leu Lys 685 690 695 700	2294
agt gtg tca aac agt att ctt gaa tgt tat acc cca gcc caa acc att Ser Val Ser Asn Ser Ile Leu Glu Cys Tyr Thr Pro Ala Gln Thr Ile 705 710 715	2342
tca act gag ttt gct gtt aaa ttg aaa att gac tta gcc aac cga gag Ser Thr Glu Phe Ala Val Lys Leu Lys Ile Asp Leu Ala Asn Arg Glu 720 725 730	2390
aca agc atc ttc agt tac cgt gaa gat ccc att gtc tat gaa att cat Thr Ser Ile Phe Ser Tyr Arg Glu Asp Pro Ile Val Tyr Glu Ile His 735 740 745	2438
cca acc aaa tct ttt att agt ggt ggg agc aca ata aca ggt gtt ggg Pro Thr Lys Ser Phe Ile Ser Gly Gly Ser Thr Ile Thr Gly Val Gly 750 755 760	2486
aaa aac ctg aat tca gtt agt gtc ccg aga atg gtc ata aat gtg cat Lys Asn Leu Asn Ser Val Ser Val Pro Arg Met Val Ile Asn Val His	2534

gag ata atc tgt tgt acc act cct tcc ctg caa cag ctg aat ctg caa	2630
Glu Ile Ile Cys Cys Thr Thr Pro Ser Leu Gln Gln Leu Asn Leu Gln	
800 805 810	
ctc ccc ctg aaa acc aaa gcc ttt ttc atg tta gat ggg atc ctt tcc	2678
Leu Pro Leu Lys Thr Lys Ala Phe Phe Met Leu Asp Gly Ile Leu Ser	
815 820 825	
aaa tac ttt gat ctc att tat gta cat aat cct gtg ttt aag cct ttt	2726
Lys Tyr Phe Asp Leu Ile Tyr Val His Asn Pro Val Phe Lys Pro Phe	
830 835 840	
gaa aag cca gtg atg atc tca atg gcc aat gaa aat gta ctg gaa att	2774
Glu Lys Pro Val Met Ile Ser Met Gly Asn Glu Asn Val Leu Glu Ile	
845 850 855 860	
aag gga aat gat att gac cct gaa gca gtt aaa ggt gaa gtg tta aaa	2822
Lys Gly Asn Asp Ile Asp Pro Glu Ala Val Lys Gly Glu Val Leu Lys	
865 870 875	
gtt gga aat aag agc tgt gag aat ata cac tta cat tct gaa gcc gtt	2870
Val Gly Asn Lys Ser Cys Glu Asn Ile His Leu His Ser Glu Ala Val	
880 885 890	
tta tgc acg gtc ccc aat gac ctg ctg aaa ttg aac agc gag cta aat	2918
Leu Cys Thr Val Pro Asn Asp Leu Leu Lys Leu Asn Ser Glu Leu Asn	
895 900 905	
ata gag tgg aag caa gca att tct tca acc gtc ctt gga aaa gta ata	2966
Ile Glu Trp Lys Gln Ala Ile Ser Ser Thr Val Leu Gly Lys Val Ile	
910 915 920	
gtt caa cca gat cag aat ttc aca gga ttg att gct ggt gtt gtc tca	3014
Val Gln Pro Asp Gln Asn Phe Thr Gly Leu Ile Ala Gly Val Val Ser	
925 930 935 940	
ata tca aca gca ctg tta tta cta ctt ggg ttt ttc ctg tgg ctg aaa	3062
Ile Ser Thr Ala Leu Leu Leu Leu Gly Phe Phe Leu Trp Leu Lys	
945 950 955	
aag aga aag caa att aaa gat ctg gcc agt gaa tta gtt cgc tac gat	3110
Lys Arg Lys Lys Gln Ile Lys Asp Leu Gly Ser Glu Leu Val Arg Tyr Asp	
960 965 970	
gca aga gta cac act cct cat ttg gat agg ctt gta agt gcc cga agt	3158
Ala Arg Val His Thr Pro His Leu Asp Arg Leu Val Ser Ala Arg Ser	
975 980 985	
gta agc cca act aca gaa atg gtt tca aat gaa tct gta gac tac cga	3206
Val Ser Pro Thr Thr Glu Met Val Ser Asn Glu Ser Val Asp Tyr Arg	
990 995 1000	



Ser	Cys	Arg	Gln	Val	Gln	Tyr	Pro	Leu	Thr	Asp	Met	Ser	Pro	Ile		
1020					1025					1030						
cta	act	agt	ggg	gac	tct	gat	ata	tcc	agt	cca	tta	ctg	caa	aat	3341	
Leu	Thr	Ser	Gly	Asp	Ser	Asp	Ile	Ser	Ser	Pro	Leu	Leu	Gln	Asn		
1035					1040					1045						
act	gtc	cac	att	gac	ctc	agt	gct	cta	aat	cca	gag	ctg	gtc	cag	3385	
Thr	Val	His	Ile	Asp	Leu	Ser	Ala	Leu	Asn	Pro	Glu	Leu	Val	Gln		
1050					1055					1060						
gca	gtg	cag	cat	gta	gtg	att	ggg	ccc	agt	agc	ctg	att	gtg	cat	3431	
Ala	Val	Gln	His	Val	Val	Ile	Gly	Pro	Ser	Ser	Leu	Ile	Val	His		
1065					1070					1075						
ttc	aat	gaa	gtc	ata	gga	aga	ggg	cat	ttt	ggc	tgt	gta	tat	cat	3476	
Phe	Asn	Glu	Val	Ile	Gly	Arg	Gly	His	Phe	Gly	Cys	Val	Tyr	His		
1080					1085					1090						
ggg	act	ttg	ttg	gac	aat	gat	ggc	aag	aaa	att	cac	tgt	gct	gtg	3521	
Gly	Thr	Leu	Leu	Asp	Asn	Asp	Gly	Lys	Lys	Ile	His	Cys	Ala	Val		
1095					1100					1105						
aaa	tcc	ttg	aac	aga	atc	act	gac	ata	gga	gaa	gtt	tcc	caa	ttt	3566	
Lys	Ser	Leu	Asn	Arg	Ile	Thr	Asp	Ile	Gly	Glu	Val	Ser	Gln	Phe		
1110					1115					1120						
ctg	acc	gag	gga	atc	atc	atg	aaa	gat	ttt	agt	cat	ccc	aat	gtc	3611	
Leu	Thr	Glu	Gly	Ile	Ile	Met	Lys	Asp	Phe	Ser	His	Pro	Asn	Val		
1125					1130					1135						
ctc	tgc	ctc	ctg	gga	atc	tgc	ctg	cga	agt	gaa	ggg	tct	ccg	ctg	3656	
Leu	Ser	Leu	Leu	Gly	Ile	Cys	Leu	Arg	Ser	Glu	Gly	Ser	Pro	Leu		
1140					1145					1150						
gtg	gtc	cta	cca	tac	atg	aaa	cat	gga	gat	ctt	cga	aat	ttc	att	3701	
Val	Val	Leu	Pro	Tyr	Met	Lys	His	Gly	Asp	Leu	Arg	Asn	Phe	Ile		
1155					1160					1165						
cga	aat	gag	act	cat	aat	cca	act	gta	aaa	gat	ctt	att	ggc	ttt	3746	
Arg	Asn	Glu	Thr	His	Asn	Pro	Thr	Val	Lys	Asp	Leu	Ile	Gly	Phe		
1170					1175					1180						
ggc	ctt	caa	gta	gcc	aaa	ggc	atg	aaa	tat	ctt	gca	agc	aaa	aag	3791	
Gly	Leu	Gln	Val	Ala	Lys	Gly	Met	Lys	Tyr	Leu	Ala	Ser	Lys	Lys		
1185					1190					1195						
ttt	gtc	cac	aga	gac	ttg	gct	gca	aga	aac	tgt	atg	ctg	gat	gaa	3836	
Phe	Val	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Cys	Met	Leu	Asp	Glu		
1200					1205					1210						

Tyr Arg Lys His Tyr Tyr Ser Val His Asn Lys Thr Gly Ala Lys

1230	1235	1240	
ctg cca gtg aag tgg atg gct ttg gaa agt	ctg caa act caa aag	3971	
Leu Pro Val Lys Trp Met Ala Leu Glu Ser	Leu Gln Thr Gln Lys		
1245	1250	1255	
ttt acc acc aag tca gat gtg tgg tcc ttt ggc	gtc gtc ctc tgg	4016	
Phe Thr Thr Lys Ser Asp Val Trp Ser Phe	Gly Val Val Leu Trp		
1260	1265	1270	
gag ctg atg aca aga gga gcc cca cct tat cct	gac gta aac acc	4061	
Glu Leu Met Thr Arg Gly Ala Pro Pro Tyr Pro	Asp Val Asn Thr		
1275	1280	1285	
ttt gat ata act gtt tac ttg ttg caa ggg aga	aga ctc cta caa	4106	
Phe Asp Ile Thr Val Tyr Leu Leu Gln Gly Arg	Arg Leu Leu Gln		
1290	1295	1300	
ccc gaa tac tgc cca gac ccc tta tat gaa gta	atg cta aaa tgc	4151	
Pro Glu Tyr Cys Pro Asp Pro Leu Tyr Glu Val	Met Leu Lys Cys		
1305	1310	1315	
tgg cac cct aaa gcc gaa atg cgc cca tcc ttt	tct gaa ctg gtg	4196	
Trp His Pro Lys Ala Glu Met Arg Pro Ser Phe	Ser Glu Leu Val		
1320	1325	1330	
tcc cgg ata tca gcg atc ttc tct act ttc att	ggg gag cac tat	4241	
Ser Arg Ile Ser Ala Ile Phe Ser Thr Phe Ile	Gly Glu His Tyr		
1335	1340	1345	
gtc cat gtg aac gct act tat gtg aac gta aaa	tgt gtc gct ccg	4286	
Val His Val Asn Ala Thr Tyr Val Asn Val Lys	Cys Val Ala Pro		
1350	1355	1360	
tat cct tct ctg ttg tca tca gaa gat aac gct	gat gat gag gtg	4331	
Tyr Pro Ser Leu Leu Ser Ser Glu Asp Asn Ala	Asp Asp Glu Val		
1365	1370	1375	
gac aca cga cca gcc tcc ttc tgg gag aca tca	tag tgctagtact	4377	
Asp Thr Arg Pro Ala Ser Phe Trp Glu Thr Ser			
1380	1385	1390	
atgtcaaaagc aacagtcac acccttgcca atgggttttt	cactgcccga cctttaaaag	4437	
gccatcgata ttctttgtct cttgcacaaat tgcactatta	ataggacttg tattgttatt	4497	
taaattactg gattctaagg aattttcttat ctgacagagc	atcagaacca gaggcttggt	4557	
cccacaggcc agggaccaat gcgctgcag		4586	

<221> SIGNAL  
<222> (1)..(24)  
<223>

<220>  
<221> gene  
<222> (1)..(1390)  
<223> C-Met Proto Oncogene

<220>  
<221> CHAIN  
<222> (25)..(1390)  
<223> HEPATOCYTE GROWTH FACTOR RECEPTOR

<220>  
<221> DOMAIN  
<222> (25)..(932)  
<223> EXTRACELLULAR (POTENTIAL)

<220>  
<221> TRANSMEM  
<222> (933)..(955)  
<223> POTENTIAL

<220>  
<221> DOMAIN  
<222> (956)..(1390)  
<223> CYTOPLASMIC (POTENTIAL)

<220>  
<221> DOMAIN  
<222> (55)..(500)  
<223> SEMA.

<220>  
<221> DOMAIN  
<222> (1078)..(1345)  
<223> PROTEIN KINASE

<220>  
<221> SITE  
<222> (1009)..(1010)  
<223> Breakpoint for Translocation to Form TPR-Met Oncogene

<220>  
<221> NP\_BIND  
<222> (1084)..(1092)  
<223> ATP (BY SIMILARITY)

<220>  
<221> BINDING

... ..  
... .. 1. 4 1. 1. 4

<223> BY SIMILARITY

<220>

<221> SITE

<222> (307)..(308)

<223> CLEAVAGE (POTENTIAL)

<220>

<221> MOD\_RES

<222> (1235)..(1235)

<223> PHOSPHORYLATION (AUTO-)

<220>

<221> CARBOHYD

<222> (45)..(45)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

<221> CARBOHYD

<222> (106)..(106)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

<221> CARBOHYD

<222> (149)..(149)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

<221> CARBOHYD

<222> (202)..(202)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

<221> CARBOHYD

<222> (399)..(399)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

<221> CARBOHYD

<222> (405)..(405)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

<221> CARBOHYD

<222> (607)..(607)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>

<221> CARBOHYD

<222> (635)..(635)

<223> N-LINKED (GLCNAC...) (POTENTIAL)

<220>  
 <221> CARBOHYD  
 <222> (879)..(879)  
 <223> N-LINKED (GLCNAC...) (POTENTIAL)  
  
 <220>  
 <221> CARBOHYD  
 <222> (930)..(930)  
 <223> N-LINKED (GLCNAC...) (POTENTIAL)  
  
 <220>  
 <221> VARSPLIC  
 <222> (755)..(755)  
 <223> S -> STWWKEPLNIVSFLFCFAS (IN ISOFORM 2)  
  
 <220>  
 <221> VARIANT  
 <222> (320)..(320)  
 <223> A -> V  
  
 <220>  
 <221> VARIANT  
 <222> (1131)..(1131)  
 <223> M -> T (IN HPRC; GERMLINE MUTATION)  
  
 <220>  
 <221> VARIANT  
 <222> (1188)..(1188)  
 <223> V -> L (IN HPRC; GERMLINE MUTATION)  
  
 <220>  
 <221> VARIANT  
 <222> (1195)..(1195)  
 <223> L -> V (IN HPRC; GERMLINE MUTATION)  
  
 <220>  
 <221> VARIANT  
 <222> (1220)..(1220)  
 <223> V -> I (IN HPRC; GERMLINE MUTATION)  
  
 <220>  
 <221> VARIANT  
 <222> (1228)..(1228)  
 <223> D -> N (IN HPRC; GERMLINE MUTATION)  
  
 <220>  
 <221> VARIANT  
 <222> (1228)..(1228)  
 <223> D -> H (IN HPRC; GERMLINE MUTATION)  
  
 <220>

<220>  
 <221> VARIANT

<222> (1230)..(1230)  
<223> Y -> H (IN HPRC; GERMLINE MUTATION)

<220>  
<221> VARIANT  
<222> (1250)..(1250)  
<223> M -> T (IN HPRC; GERMLINE MUTATION)

<400> 12

Met Lys Ala Pro Ala Val Leu Ala Pro Gly Ile Leu Val Leu Leu Phe  
1 5 10 15

Thr Leu Val Gln Arg Ser Asn Gly Glu Cys Lys Glu Ala Leu Ala Lys  
20 25 30

Ser Glu Met Asn Val Asn Met Lys Tyr Gln Leu Pro Asn Phe Thr Ala  
35 40 45

Glu Thr Pro Ile Gln Asn Val Ile Leu His Glu His His Ile Phe Leu  
50 55 60

Gly Ala Thr Asn Tyr Ile Tyr Val Leu Asn Glu Glu Asp Leu Gln Lys  
65 70 75 80

Val Ala Glu Tyr Lys Thr Gly Pro Val Leu Glu His Pro Asp Cys Phe  
85 90 95

Pro Cys Gln Asp Cys Ser Ser Lys Ala Asn Leu Ser Gly Gly Val Trp  
100 105 110

Lys Asp Asn Ile Asn Met Ala Leu Val Val Asp Thr Tyr Tyr Asp Asp  
115 120 125

Gln Leu Ile Ser Cys Gly Ser Val Asn Arg Gly Thr Cys Gln Arg His  
130 135 140

Val Phe Pro His Asn His Thr Ala Asp Ile Gln Ser Glu Val His Cys  
145 150 155 160

1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250

Ile Asn Phe Phe Val Gly Asn Thr Ile Asn Ser Ser Tyr Phe Pro Asp  
195 200 205

His Pro Leu His Ser Ile Ser Val Arg Arg Leu Lys Glu Thr Lys Asp  
210 215 220

Gly Phe Met Phe Leu Thr Asp Gln Ser Tyr Ile Asp Val Leu Pro Glu  
225 230 235 240

Phe Arg Asp Ser Tyr Pro Ile Lys Tyr Val His Ala Phe Glu Ser Asn  
245 250 255

Asn Phe Ile Tyr Phe Leu Thr Val Gln Arg Glu Thr Leu Asp Ala Gln  
260 265 270

Thr Phe His Thr Arg Ile Ile Arg Phe Cys Ser Ile Asn Ser Gly Leu  
275 280 285

His Ser Tyr Met Glu Met Pro Leu Glu Cys Ile Leu Thr Glu Lys Arg  
290 295 300

Lys Lys Arg Ser Thr Lys Lys Glu Val Phe Asn Ile Leu Gln Ala Ala  
305 310 315 320

Tyr Val Ser Lys Pro Gly Ala Gln Leu Ala Arg Gln Ile Gly Ala Ser  
325 330 335

Leu Asn Asp Asp Ile Leu Phe Gly Val Phe Ala Gln Ser Lys Pro Asp  
340 345 350

Ser Ala Glu Pro Met Asp Arg Ser Ala Met Cys Ala Phe Pro Ile Lys  
355 360 365

Tyr Val Asn Asp Phe Phe Asn Lys Ile Val Asn Lys Asn Asn Val Arg  
370 375 380

Cys Leu Gln His Phe Tyr Gly Pro Asn His Glu His Cys Phe Asn Arg

Arg Thr Glu Phe Thr Thr Ala Leu Gln Arg Val Asp Leu Phe Met Gly  
420 425 430

Gln Phe Ser Glu Val Leu Leu Thr Ser Ile Ser Thr Phe Ile Lys Gly  
435 440 445

Asp Leu Thr Ile Ala Asn Leu Gly Thr Ser Glu Gly Arg Phe Met Gln  
450 455 460

Val Val Val Ser Arg Ser Gly Pro Ser Thr Pro His Val Asn Phe Leu  
465 470 475 480

Leu Asp Ser His Pro Val Ser Pro Glu Val Ile Val Glu His Thr Leu  
485 490 495

Asn Gln Asn Gly Tyr Thr Leu Val Ile Thr Gly Lys Lys Ile Thr Lys  
500 505 510

Ile Pro Leu Asn Gly Leu Gly Cys Arg His Phe Gln Ser Cys Ser Gln  
515 520 525

Cys Leu Ser Ala Pro Pro Phe Val Gln Cys Gly Trp Cys His Asp Lys  
530 535 540

Cys Val Arg Ser Glu Glu Cys Leu Ser Gly Thr Trp Thr Gln Gln Ile  
545 550 555 560

Cys Leu Pro Ala Ile Tyr Lys Val Phe Pro Asn Ser Ala Pro Leu Glu  
565 570 575

Gly Gly Thr Arg Leu Thr Ile Cys Gly Thr Asp Phe Gly Ile Arg Arg  
580 585 590

Asn Asn Lys Phe Asp Leu Lys Lys Thr Arg Val Leu Leu Gly Asn Glu  
595 600 605

Ser Cys Thr Leu Thr Leu Ser His Ser Thr Met Asn Thr Leu Lys Cys  
610 615 620





865

870

875

880

Ser Cys Glu Asn Ile His Leu His Ser Glu Ala Val Leu Cys Thr Val  
 885 890 895

Pro Asn Asp Leu Leu Lys Leu Asn Ser Glu Leu Asn Ile Glu Trp Lys  
 900 905 910

Gln Ala Ile Ser Ser Thr Val Leu Gly Lys Val Ile Val Gln Pro Asp  
 915 920 925

Gln Asn Phe Thr Gly Leu Ile Ala Gly Val Val Ser Ile Ser Thr Ala  
 930 935 940

Leu Leu Leu Leu Leu Gly Phe Phe Leu Trp Leu Lys Lys Arg Lys Gln  
 945 950 955 960

Ile Lys Asp Leu Gly Ser Glu Leu Val Arg Tyr Asp Ala Arg Val His  
 965 970 975

Thr Pro His Leu Asp Arg Leu Val Ser Ala Arg Ser Val Ser Pro Thr  
 980 985 990

Thr Glu Met Val Ser Asn Glu Ser Val Asp Tyr Arg Ala Thr Phe Pro  
 995 1000 1005

Glu Asp Gln Phe Pro Asn Ser Ser Gln Asn Gly Ser Cys Arg Gln  
 1010 1015 1020

Val Gln Tyr Pro Leu Thr Asp Met Ser Pro Ile Leu Thr Ser Gly  
 1025 1030 1035

Asp Ser Asp Ile Ser Ser Pro Leu Leu Gln Asn Thr Val His Ile  
 1040 1045 1050

Asp Leu Ser Ala Leu Asn Pro Glu Leu Val Gln Ala Val Gln His  
 1055 1060 1065

Ile Gly Arg Gly His Ile Gly Tyr Val Tyr His Gly Thr Leu Leu

1085		1090		1095
Asp Asn Asp Gly Lys Lys Ile His Cys Ala Val Lys Ser Leu Asn				
1100		1105		1110
Arg Ile Thr Asp Ile Gly Glu Val Ser Gln Phe Leu Thr Glu Gly				
1115		1120		1125
Ile Ile Met Lys Asp Phe Ser His Pro Asn Val Leu Ser Leu Leu				
1130		1135		1140
Gly Ile Cys Leu Arg Ser Glu Gly Ser Pro Leu Val Val Leu Pro				
1145		1150		1155
Tyr Met Lys His Gly Asp Leu Arg Asn Phe Ile Arg Asn Glu Thr				
1160		1165		1170
His Asn Pro Thr Val Lys Asp Leu Ile Gly Phe Gly Leu Gln Val				
1175		1180		1185
Ala Lys Gly Met Lys Tyr Leu Ala Ser Lys Lys Phe Val His Arg				
1190		1195		1200
Asp Leu Ala Ala Arg Asn Cys Met Leu Asp Glu Lys Phe Thr Val				
1205		1210		1215
Lys Val Ala Asp Phe Gly Leu Ala Arg Asp Met Tyr Asp Lys Glu				
1220		1225		1230
Tyr Tyr Ser Val His Asn Lys Thr Gly Ala Lys Leu Pro Val Lys				
1235		1240		1245
Trp Met Ala Leu Glu Ser Leu Gln Thr Gln Lys Phe Thr Thr Lys				
1250		1255		1260
Ser Asp Val Trp Ser Phe Gly Val Val Leu Trp Glu Leu Met Thr				
1265		1270		1275

Pro Asp Pro Leu Tyr Glu Val Met Leu Lys Cys Trp His Pro Lys  
1310 1315 1320

Ala Glu Met Arg Pro Ser Phe Ser Glu Leu Val Ser Arg Ile Ser  
1325 1330 1335

Ala Ile Phe Ser Thr Phe Ile Gly Glu His Tyr Val His Val Asn  
1340 1345 1350

Ala Thr Tyr Val Asn Val Lys Cys Val Ala Pro Tyr Pro Ser Leu  
1355 1360 1365

Leu Ser Ser Glu Asp Asn Ala Asp Asp Glu Val Asp Thr Arg Pro  
1370 1375 1380

Ala Ser Phe Trp Glu Thr Ser  
1385 1390

<210> 13

<211> 2558

<212> DNA

<213> Homo sapiens

<220>

<221> gene

<222> (1)..(2558)

<223> Chondroitin sulfate proteoglycan BEHAB/brevican mRNA, GPI isoform  
, complete cds

<220>

<221> CDS

<222> (48)..(2063)

<223>

<300>

<301> Gary,S.C., Zerillo,C.A., Chiang,V.L., Gaw,J.U., Gray,G. and  
Hockfield,S.

<302> cDNA cloning, chromosomal localization, and expression analysis of  
human BEHAB/brevican, a brain specific proteoglycan regulated during cortical  
development and in glioma

<303> Gene

<304> 256

<305> 1

<400> 13

tgtggcactg cctgcgtacc caaccccagc cctgggtagc ctgcagc atg gcc cag	56
Met Ala Gln	
1	
ctg ttc ctg ccc ctg ctg gca gcc ctg gtc ctg gcc cag gct cct gca	104
Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala Gln Ala Pro Ala	
5 10 15	
gct tta gca gat gtt ctg gaa gga gac agc tca gag gac cgc gct ttt	152
Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser Glu Asp Arg Ala Phe	
20 25 30 35	
cgc gtg cgc atc gcg ggc gac gcg cca ctg cag ggc gtg ctc ggc ggc	200
Arg Val Arg Ile Ala Gly Asp Ala Pro Leu Gln Gly Val Leu Gly Gly	
40 45 50	
gcc ctc acc atc cct tgc cac gtc cac tac ctg cgg cca ccg ccg agc	248
Ala Leu Thr Ile Pro Cys His Val His Tyr Leu Arg Pro Pro Pro Ser	
55 60 65	
cgc cgg gct gtg ctg ggc tct ccg cgg gtc aag tgg act ttc ctg tcc	296
Arg Arg Ala Val Leu Gly Ser Pro Arg Val Lys Trp Thr Phe Leu Ser	
70 75 80	
cgg ggc cgg gag gca gag gtg ctg gtg gcg cgg gga gtg cgc gtc aag	344
Arg Gly Arg Glu Ala Glu Val Leu Val Ala Arg Gly Val Arg Val Lys	
85 90 95	
gtg aac gag gcc tac cgg ttc cgc gtg gca ctg cct gcg tac cca gcg	392
Val Asn Glu Ala Tyr Arg Phe Arg Val Ala Leu Pro Ala Tyr Pro Ala	
100 105 110 115	
tgc ctc acc gac gtc tcc ctg gcg ctg agc gag ctg cgc ccc aac gac	440
Ser Leu Thr Asp Val Ser Leu Ala Leu Ser Glu Leu Arg Pro Asn Asp	
120 125 130	
tca ggt atc tat cgc tgt gag gtc cag cac ggc atc gat gac agc agc	488
Ser Gly Ile Tyr Arg Cys Glu Val Gln His Gly Ile Asp Asp Ser Ser	
135 140 145	
gac gct gtg gag gtc aag gtc aaa ggg gtc gtc ttt ctc tac cga gag	536
Asp Ala Val Glu Val Lys Val Lys Gly Val Val Phe Leu Tyr Arg Glu	
150 155 160	
ggc tct gcc cgc tat gct ttc tcc ttt tct ggg gcc cag gag gcc tgt	584
Gly Ser Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln Glu Ala Cys	
165 170 175	
gcc cgc att gga gcc cac atc gcc acc ccg gag cag ctc tat gcc gcc	632
Ala Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu Tyr Ala Ala	

acc gtg agg tat ccc atc cag acc cca cga gag gcc tgt tac gga gac	728
Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys Tyr Gly Asp	
215 220 225	
atg gat ggc ttc ccc ggg gtc cgg aac tat ggt gtg gtg gac ccg gat	776
Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val Asp Pro Asp	
230 235 240	
gac ctc tat gat gtg tac tgt tat gct gaa gac cta aat gga gaa ttg	824
Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu Asn Gly Glu Leu	
245 250 255	
ttc ctg ggt gac cct cca gag aag ctg aca ttg gag gaa gca cgg gcg	872
Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu Glu Glu Ala Arg Ala	
260 265 270 275	
tac tgc cag gag cgg ggt gca gag att gcc acc acg ggc caa ctg tat	920
Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala Thr Thr Gly Gln Leu Tyr	
280 285 290	
gca gcc tgg gat ggt ggc ctg gac cac tgc agc cca ggg tgg cta gct	968
Ala Ala Trp Asp Gly Gly Leu Asp His Cys Ser Pro Gly Trp Leu Ala	
295 300 305	
gat ggc agt gtg cgc tac ccc atc gtc aca ccc agc cag cgc tgt ggt	1016
Asp Gly Ser Val Arg Tyr Pro Ile Val Thr Pro Ser Gln Arg Cys Gly	
310 315 320	
ggg ggc ttg cct ggt gtc aag act ctc ttc ctc ttc ccc aac cag act	1064
Gly Gly Leu Pro Gly Val Lys Thr Leu Phe Leu Phe Pro Asn Gln Thr	
325 330 335	
ggc ttc ccc aat aag cac agc cgc ttc aac gtc tac tgc ttc cga gac	1112
Gly Phe Pro Asn Lys His Ser Arg Phe Asn Val Tyr Cys Phe Arg Asp	
340 345 350 355	
tcg gcc cag cct tct gcc atc cct gag gcc tcc aac cca gcc tcc aac	1160
Ser Ala Gln Pro Ser Ala Ile Pro Glu Ala Ser Asn Pro Ala Ser Asn	
360 365 370	
cca gcc tct gat gga cta gag gct atc gtc aca gtg aca gag acc ctg	1208
Pro Ala Ser Asp Gly Leu Glu Ala Ile Val Thr Val Thr Glu Thr Leu	
375 380 385	
gag gaa ctg cag ctg cct cag gaa gcc aca gag agt gaa tcc cgt ggg	1256
Glu Glu Leu Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu Ser Arg Gly	
390 395 400	
gcc atc tac tcc atc ccc atc atg gag gac gga gga ggt gga agc tcc	1304
Ala Ile Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly Gly Ser Ser	
405 410 415	

Thr	Gln	Ser	Met	Val	Pro	Pro	Thr	Gly	Phe	Ser	Glu	Glu	Glu	Gly	Lys	
				440					445					450		
gca	ttg	gag	gaa	gaa	gag	aaa	tat	gaa	gat	gaa	gaa	gag	aaa	gag	gag	1448
Ala	Leu	Glu	Glu	Glu	Glu	Lys	Tyr	Glu	Asp	Glu	Glu	Glu	Lys	Glu	Glu	
			455					460					465			
gaa	gaa	gaa	gag	gag	gag	gtg	gag	gat	gag	gct	ctg	tgg	gca	tgg	ccc	1496
Glu	Glu	Glu	Glu	Glu	Glu	Val	Glu	Asp	Glu	Ala	Leu	Trp	Ala	Trp	Pro	
			470				475					480				
agc	gag	ctc	agc	agc	ccg	ggc	cct	gag	gcc	tct	ctc	ccc	act	gag	cca	1544
Ser	Glu	Leu	Ser	Ser	Pro	Gly	Pro	Glu	Ala	Ser	Leu	Pro	Thr	Glu	Pro	
	485					490					495					
gca	gcc	cag	gag	gag	tca	ctc	tcc	cag	gcg	cca	gca	agg	gca	gtc	ctg	1592
Ala	Ala	Gln	Glu	Glu	Ser	Leu	Ser	Gln	Ala	Pro	Ala	Arg	Ala	Val	Leu	
500					505					510					515	
cag	cct	ggt	gca	tca	cca	ctt	cct	gat	gga	gag	tca	gaa	gct	tcc	agg	1640
Gln	Pro	Gly	Ala	Ser	Pro	Leu	Pro	Asp	Gly	Glu	Ser	Glu	Ala	Ser	Arg	
			520						525				530			
cct	cca	agg	gtc	cat	gga	cca	cct	act	gag	act	ctg	ccc	act	ccc	agg	1688
Pro	Pro	Arg	Val	His	Gly	Pro	Pro	Thr	Glu	Thr	Leu	Pro	Thr	Pro	Arg	
			535					540					545			
gag	agg	aac	cta	gca	tcc	cca	tca	cct	tcc	act	ctg	gtt	gag	gca	aga	1736
Glu	Arg	Asn	Leu	Ala	Ser	Pro	Ser	Pro	Ser	Thr	Leu	Val	Glu	Ala	Arg	
	550						555				560					
gag	gtg	ggg	gag	gca	act	ggt	ggt	cct	gag	cta	tct	ggg	gtc	cct	cga	1784
Glu	Val	Gly	Glu	Ala	Thr	Gly	Gly	Pro	Glu	Leu	Ser	Gly	Val	Pro	Arg	
	565					570					575					
gga	gag	agc	gag	gag	aca	gga	agc	tcc	gag	ggt	gcc	cct	tcc	ctg	ctt	1832
Gly	Glu	Ser	Glu	Glu	Thr	Gly	Ser	Ser	Glu	Gly	Ala	Pro	Ser	Leu	Leu	
580					585				590					595		
cca	gcc	aca	cgg	gcc	cct	gag	ggt	acc	agg	gag	ctg	gag	gcc	ccc	tct	1880
Pro	Ala	Thr	Arg	Ala	Pro	Glu	Gly	Thr	Arg	Glu	Leu	Glu	Ala	Pro	Ser	
			600					605					610			
gaa	gat	aat	tct	gga	aga	act	gcc	cca	gca	ggg	acc	tca	gtg	cag	gcc	1928
Glu	Asp	Asn	Ser	Gly	Arg	Thr	Ala	Pro	Ala	Gly	Thr	Ser	Val	Gln	Ala	
			615				620					625				
cag	cca	gtg	ctg	ccc	act	gac	agc	gcc	agc	cga	ggt	gga	gtg	gcc	gtg	1976
Gln	Pro	Val	Leu	Pro	Thr	Asp	Ser	Ala	Ser	Arg	Gly	Gly	Val	Ala	Val	
	630					635					640					

Leu Leu Leu Phe Phe Phe Leu Gln Leu Thr Val Thr

660

665

670

```

tttaaccac catcatccca aactctctg tcttttgct tcattctctt acccactct 2133
acctatgggt ctccaatctc ggatatacac cttgtgggta tctcagctct ccgcgtcttt 2193
acctgtgat ccagccccg ccactgacaa tctgtgacct ttcctgcca ttgggccctc 2253
cacctgtggc tcacatctcg ccagccccac agagcactct caggctctc caagggctct 2313
catcacctat tgcagcttc agggctcggc ctattttcca ctactcctt catccgctg 2373
tgtgcgctcc cctttagctg cctcctattg atctcagga agcctgggag tcccttctca 2433
ccctcaaac tccggagtcc aggagaacc gtacccccac agagccttaa gcaactactt 2493
ctgtgaagta ttttttgact gtttcatgga aaacaagcct tggaaataaa tctctattaa 2553
accgc 2558

```

&lt;210&gt; 14

&lt;211&gt; 671

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; gene

&lt;222&gt; (1)..(671)

&lt;223&gt; Chondroitin sulfate proteoglycan BEHAB/brevican

&lt;400&gt; 14

```

Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala Gln
1           5           10           15

```

```

Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser Glu Asp
20           25           30

```

```

Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu Gln Gly Val
35           40           45

```

```

Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His Tyr Leu Arg Pro
50           55           60

```

```

Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro Arg Val Lys Thr Thr

```



Arg Val Lys Val Asn Glu Ala Tyr Arg Phe Arg Val Ala Leu Pro Ala  
100 105 110

Tyr Pro Ala Ser Leu Thr Asp Val Ser Leu Ala Leu Ser Glu Leu Arg  
115 120 125

Pro Asn Asp Ser Gly Ile Tyr Arg Cys Glu Val Gln His Gly Ile Asp  
130 135 140

Asp Ser Ser Asp Ala Val Glu Val Lys Val Lys Gly Val Val Phe Leu  
145 150 155 160

Tyr Arg Glu Gly Ser Ala Arg Tyr Ala Phe Ser Phe Ser Gly Ala Gln  
165 170 175

Glu Ala Cys Ala Arg Ile Gly Ala His Ile Ala Thr Pro Glu Gln Leu  
180 185 190

Tyr Ala Ala Tyr Leu Gly Gly Tyr Glu Gln Cys Asp Ala Gly Trp Leu  
195 200 205

Ser Asp Gln Thr Val Arg Tyr Pro Ile Gln Thr Pro Arg Glu Ala Cys  
210 215 220

Tyr Gly Asp Met Asp Gly Phe Pro Gly Val Arg Asn Tyr Gly Val Val  
225 230 235 240

Asp Pro Asp Asp Leu Tyr Asp Val Tyr Cys Tyr Ala Glu Asp Leu Asn  
245 250 255

Gly Glu Leu Phe Leu Gly Asp Pro Pro Glu Lys Leu Thr Leu Glu Glu  
260 265 270

Ala Arg Ala Tyr Cys Gln Glu Arg Gly Ala Glu Ile Ala Thr Thr Gly  
275 280 285

Gln Leu Tyr Ala Ala Thr Asp Gly Gly Leu Asp His Cys Ser Pro Gly  
290 295 300

Arg Cys Gly Gly Gly Leu Pro Gly Val Lys Thr Leu Phe Leu Phe Pro  
325 330 335

Asn Gln Thr Gly Phe Pro Asn Lys His Ser Arg Phe Asn Val Tyr Cys  
340 345 350

Phe Arg Asp Ser Ala Gln Pro Ser Ala Ile Pro Glu Ala Ser Asn Pro  
355 360 365

Ala Ser Asn Pro Ala Ser Asp Gly Leu Glu Ala Ile Val Thr Val Thr  
370 375 380

Glu Thr Leu Glu Glu Leu Gln Leu Pro Gln Glu Ala Thr Glu Ser Glu  
385 390 395 400

Ser Arg Gly Ala Ile Tyr Ser Ile Pro Ile Met Glu Asp Gly Gly Gly  
405 410 415

Gly Ser Ser Thr Pro Glu Asp Pro Ala Glu Ala Pro Arg Thr Leu Leu  
420 425 430

Glu Phe Glu Thr Gln Ser Met Val Pro Pro Thr Gly Phe Ser Glu Glu  
435 440 445

Glu Gly Lys Ala Leu Glu Glu Glu Glu Lys Tyr Glu Asp Glu Glu Glu  
450 455 460

Lys Glu Glu Glu Glu Glu Glu Glu Glu Val Glu Asp Glu Ala Leu Trp  
465 470 475 480

Ala Trp Pro Ser Glu Leu Ser Ser Pro Gly Pro Glu Ala Ser Leu Pro  
485 490 495

Thr Glu Pro Ala Ala Gln Glu Glu Ser Leu Ser Gln Ala Pro Ala Arg  
500 505 510

Ala Val Leu Gln Pro Gly Ala Ser Pro Leu Pro Asp Gly Glu Ser Glu  
515 520 525

Thr Thr Arg Thr Arg Arg Leu Ala Thr Thr Thr Thr Thr Thr Thr Thr

545 550 555 560

Glu Ala Arg Glu Val Gly Glu Ala Thr Gly Gly Pro Glu Leu Ser Gly  
565 570 575

Val Pro Arg Gly Glu Ser Glu Glu Thr Gly Ser Ser Glu Gly Ala Pro  
580 585 590

Ser Leu Leu Pro Ala Thr Arg Ala Pro Glu Gly Thr Arg Glu Leu Glu  
595 600 605

Ala Pro Ser Glu Asp Asn Ser Gly Arg Thr Ala Pro Ala Gly Thr Ser  
610 615 620

Val Gln Ala Gln Pro Val Leu Pro Thr Asp Ser Ala Ser Arg Gly Gly  
625 630 635 640

Val Ala Val Val Pro Ala Ser Gly Asn Ser Ala Gln Gly Ser Thr Ala  
645 650 655

Leu Ser Ile Leu Leu Leu Phe Phe Pro Leu Gln Leu Trp Val Thr  
660 665 670

<210> 15  
<211> 2316  
<212> DNA  
<213> Homo sapiens

<220>  
<221> gene  
<222> (1)..(2316)  
<223> Human mRNA for CD44E (epithelial form)

<220>  
<221> CDS  
<222> (124)..(1605)  
<223>

<300>  
<301> Stamenkovic,I., Aruffo,A., Amiot,M. and Seed,B.  
<302> The hematopoietic and epithelial forms of CD44 are distinct  
polypeptides with different adhesion potentials for hyaluronate-bearing cells  
<303> EMBO J.



195	200	205	
gac agt ccc tgg atc acc gac agc aca gac aga atc cct cgt acc aat			792
Asp Ser Pro Trp Ile Thr Asp Ser Thr Asp Arg Ile Pro Arg Thr Asn			
210	215	220	
atg gac tcc agt cat agt aca acg ctt cag cct act gca aat cca aac			840
Met Asp Ser Ser His Ser Thr Thr Leu Gln Pro Thr Ala Asn Pro Asn			
225	230	235	
aca ggt ttg gtg gaa gat ttg gac agg aca gga cct ctt tca atg aca			888
Thr Gly Leu Val Glu Asp Leu Asp Arg Thr Gly Pro Leu Ser Met Thr			
240	245	250	255
acg cag cag agt aat tct cag agc ttc tct aca tca cat gaa ggc ttg			936
Thr Gln Gln Ser Asn Ser Gln Ser Phe Ser Thr Ser His Glu Gly Leu			
260	265	270	
gaa gaa gat aaa gac cat cca aca act tct act ctg aca tca agc aat			984
Glu Glu Asp Lys Asp His Pro Thr Thr Ser Thr Leu Thr Ser Ser Asn			
275	280	285	
agg aat gat gtc aca ggt gga aga aga gac cca aat cat tct gaa ggc			1032
Arg Asn Asp Val Thr Gly Gly Arg Arg Asp Pro Asn His Ser Glu Gly			
290	295	300	
tca act cat tta ctg gaa ggt tat acc tct cat tac cca cac acg aag			1080
Ser Thr His Leu Leu Glu Gly Tyr Thr Ser His Tyr Pro His Thr Lys			
305	310	315	
gaa agc agg acc ttc atc cca gtg acc tca gct aag act ggg tcc ttt			1128
Glu Ser Arg Thr Phe Ile Pro Val Thr Ser Ala Lys Thr Gly Ser Phe			
320	325	330	335
gga gtt act gca gtt act gtt gga gat tcc aac tct aat gtc aat cgt			1176
Gly Val Thr Ala Val Thr Val Gly Asp Ser Asn Ser Asn Val Asn Arg			
340	345	350	
tcc tta tca gga gac caa gac aca ttc cac ccc agt ggg ggg tcc cat			1224
Ser Leu Ser Gly Asp Gln Asp Thr Phe His Pro Ser Gly Gly Ser His			
355	360	365	
acc act cat gga tct gaa tca gat gga cac tca cat ggg agt caa gaa			1272
Thr Thr His Gly Ser Glu Ser Asp Gly His Ser His Gly Ser Gln Glu			
370	375	380	
ggt gga gca aac aca acc tct ggt cct ata agg aca ccc caa att cca			1320
Gly Gly Ala Asn Thr Thr Ser Gly Pro Ile Arg Thr Pro Gln Ile Pro			
385	390	395	
gaa tgg ctg atc atc tta gaa tcc ctg tta acc tta act tta att ctt			1368

aag cta gtg atc aac agt ggc aat gga gct gtg gag gac aga aag cca	1464
Lys Leu Val Ile Asn Ser Gly Asn Gly Ala Val Glu Asp Arg Lys Pro	
435 440 445	
agt gga ctc aac gga gag gcc agc aag tct cag gaa atg gtg cat ttg	1512
Ser Gly Leu Asn Gly Glu Ala Ser Lys Ser Gln Glu Met Val His Leu	
450 455 460	
gtg aac aag gag tgg tca gaa act cca gac cag ttt atg aca gct gat	1560
Val Asn Lys Glu Ser Ser Glu Thr Pro Asp Gln Phe Met Thr Ala Asp	
465 470 475	
gag aca agg aac ctg cag aat gtg gac atg aag att ggg gtg taa	1605
Glu Thr Arg Asn Leu Gln Asn Val Asp Met Lys Ile Gly Val	
480 485 490	
cacctacacc attatcttgg aaagaaacaa cgttggaac ataaccatta caggggagct	1665
gggacactta acagatgcaa tgtgctactg attgtttcat ttcgaaatcta taatagcata	1725
aaattttcta ctctttttgt tttttgtgtt ttgttcttta aagtcaggtc caatttgtaa	1785
aaacagcatt gottttctgaa attagggccc aattaataat cagcaagaat tttgatcggt	1845
tcagttcccc acttgagggc ctttcatccc tcgggtgtgc tatggatggc ttctaacaaa	1905
aacctaccac atagttatct ctgatcgcca accttgcccc ccaccagcta aggacatttc	1965
cagggttaat agggcctggc cctgggagga aatttgaatg ggtcattttg cccttcatt	2025
agcctaatec ctgggcattg ctttccactg aggttggggg ttgggggtgta ctagttacac	2085
atcttcaaca gacccccctc agaaattttt cagatgcttc tgggagacac ccaaagggtg	2145
agtctattta tctgtagtaa actatttate tgtgtttttg aaatattaaa ccttgatca	2205
gtccttttat tcagtataat tttttaaagt tactttgtca gaggcacaaa aagggtttta	2265
actgattcat aataaatate tgaacttct tcgaaaaaaaa aaaaaaaaaa a	2316

<110> 16  
 <111> 742  
 <112> PRT  
 <113> Homo sapiens

<120>  
 <121> SIGNAL  
 <122> (1)..(20)  
 <123> BY SIMILARITY

<220>  
 <221> CHAIN  
 <222> (21)..(742)  
 <223> CD44 ANTIGEN  
  
 <220>  
 <221> DOMAIN  
 <222> (11)..(649)  
 <223> EXTRACELLULAR (POTENTIAL)  
  
 <220>  
 <221> TRANSMEM  
 <222> (650)..(670)  
 <223> POTENTIAL  
  
 <220>  
 <221> DOMAIN  
 <222> (671)..(742)  
 <223> CYTOPLASMIC (POTENTIAL)  
  
 <220>  
 <221> DOMAIN  
 <222> (48)..(119)  
 <223> LINK  
  
 <220>  
 <221> MOD\_RES  
 <222> (21)..(21)  
 <223> PYRROLIDONE CARBOXYLIC ACID (PROBABLE)  
  
 <220>  
 <221> DISULFID  
 <222> (53)..(118)  
 <223> BY SIMILARITY  
  
 <220>  
 <221> DISULFID  
 <222> (77)..(97)  
 <223> BY SIMILARITY  
  
 <220>  
 <221> CARBOHYD  
 <222> (15)..(25)  
 <223> N-LINKED (GLCNAC ...) (POTENTIAL)  
  
 <220>  
 <221> CARBOHYD  
 <222> (57)..(57)  
 <223> N-LINKED (GLCNAC ...) (POTENTIAL)  
  
 <220>

<220>  
 <221> CARBOHYD

```

<222> (110)..(110)
<223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>
<221> CARBOHYD
<222> (120)..(120)
<223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>
<221> CARBOHYD
<222> (350)..(350)
<223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>
<221> CARBOHYD
<222> (548)..(548)
<223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>
<221> CARBOHYD
<222> (599)..(599)
<223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>
<221> CARBOHYD
<222> (636)..(636)
<223> N-LINKED (GLCNAC ...) (POTENTIAL)

<220>
<221> VARSPLIC
<222> (23)..(29)
<223> DLNITCR -> GVGRKRS (IN ISOFORM CD44SP)

<220>
<221> VARSPLIC
<222> (30)..(742)
<223> MISSING (IN ISOFORM CD44SP)

<220>
<221> VARSPLIC
<222> (192)..(192)
<223> G -> A (IN ISOFORM WITH ALTERNATIVE SPLICE DONOR/ACCEPTOR ON EXON
      5)

<220>
<221> VARSPLIC
<222> (193)..(223)
<223> MISSING (IN ISOFORM WITH ALTERNATIVE SPLICE DONOR/ACCEPTOR ON EXO
      N 5)

<220>

...
<221> VARSPLIC

```



<222> (224)..(265)  
 <223> MISSING (IN ISOFORM WITHOUT EXON 6)  
  
 <220>  
 <221> VARSPLIC  
 <222> (223)..(223)  
 <223> T -> H (IN ISOFORMS WITHOUT EXONS 6 TO 11)  
  
 <220>  
 <221> VARSPLIC  
 <222> (224)..(472)  
 <223> MISSING (IN ISOFORMS WITHOUT EXONS 6 TO 11)  
  
 <220>  
 <221> VARSPLIC  
 <222> (223)..(223)  
 <223> T -> R (IN ISOFORMS WITHOUT EXONS 6 TO 14)  
  
 <220>  
 <221> VARSPLIC  
 <222> (224)..(604)  
 <223> MISSING (IN ISOFORMS WITHOUT EXONS 6 TO 14)  
  
 <220>  
 <221> VARSPLIC  
 <222> (266)..(273)  
 <223> MISSING (IN ISOFORM WITH ALTERNATIVE SPLICE DONOR/ACCEPTOR ON EXO  
 N 7)  
  
 <220>  
 <221> VARSPLIC  
 <222> (385)..(385)  
 <223> I -> T (IN ISOFORMS WITHOUT EXON 10)  
  
 <220>  
 <221> VARSPLIC  
 <222> (386)..(428)  
 <223> MISSING (IN ISOFORMS WITHOUT EXON 10)  
  
 <220>  
 <221> VARSPLIC  
 <222> (506)..(506)  
 <223> Q -> R (IN ISOFORMS WITHOUT EXON 13)  
  
 <220>  
 <221> VARSPLIC  
 <222> (507)..(535)  
 <223> MISSING (IN ISOFORMS WITHOUT EXON 13)  
  
 <220>  
 <221> VARSPLIC

<221> VARSPLIC  
 <222> (507)..(535)

<223> MISSING (IN ISOFORMS WITHOUT EXON 14)

<220>

<221> VARSPLIC

<222> (675)..(675)

<223> R -> S (IN ISOFORMS WITHOUT EXON 19)

<220>

<221> VARSPLIC

<222> (676)..(742)

<223> MISSING (IN ISOFORMS WITHOUT EXON 19)

<220>

<221> VARIANT

<222> (46)..(46)

<223> R -> P (IN IN(A) ANTIGEN)

<400> 16

Met Asp Lys Phe Trp Trp His Ala Ala Trp Gly Leu Cys Leu Val Pro  
1 5 10 15

Leu Ser Leu Ala Gln Ile Asp Leu Asn Ile Thr Cys Arg Phe Ala Gly  
20 25 30

Val Phe His Val Glu Lys Asn Gly Arg Tyr Ser Ile Ser Arg Thr Glu  
35 40 45

Ala Ala Asp Leu Cys Lys Ala Phe Asn Ser Thr Leu Pro Thr Met Ala  
50 55 60

Gln Met Glu Lys Ala Leu Ser Ile Gly Phe Glu Thr Cys Arg Tyr Gly  
65 70 75 80

Phe Ile Glu Gly His Val Val Ile Pro Arg Ile His Pro Asn Ser Ile  
85 90 95

Cys Ala Ala Asn Asn Thr Gly Val Tyr Ile Leu Thr Ser Asn Thr Ser  
100 105 110

Gln Tyr Asp Thr Tyr Cys Phe Asn Ala Ser Ala Pro Pro Glu Glu Asp  
115 120 125

Ile Thr Ile Val Asn Ala Arg Gly Thr Asn Tyr Val His Arg Gly His

145 150 155 160

Tyr Arg Thr Asn Pro Glu Asp Ile Tyr Pro Ser Asn Pro Thr Asp Asp  
165 170 175

Asp Val Ser Ser Gly Ser Ser Ser Glu Arg Ser Ser Thr Ser Gly Gly  
180 185 190

Tyr Ile Phe Tyr Thr Phe Ser Thr Val His Pro Ile Pro Asp Glu Asp  
195 200 205

Ser Pro Trp Ile Thr Asp Ser Thr Asp Arg Ile Pro Ala Thr Thr Leu  
210 215 220

Met Ser Thr Ser Ala Thr Ala Thr Glu Thr Ala Thr Lys Arg Gln Glu  
225 230 235 240

Thr Trp Asp Trp Phe Ser Trp Leu Phe Leu Pro Ser Glu Ser Lys Asn  
245 250 255

His Leu His Thr Thr Thr Gln Met Ala Gly Thr Ser Ser Asn Thr Ile  
260 265 270

Ser Ala Gly Trp Glu Pro Asn Glu Glu Asn Glu Asp Glu Arg Asp Arg  
275 280 285

His Leu Ser Phe Ser Gly Ser Gly Ile Asp Asp Asp Glu Asp Phe Ile  
290 295 300

Ser Ser Thr Ile Ser Thr Thr Pro Arg Ala Phe Asp His Thr Lys Gln  
305 310 315 320

Asn Gln Asp Trp Thr Gln Trp Asn Pro Ser His Ser Asn Pro Glu Val  
325 330 335

Leu Leu Gln Thr Thr Thr Arg Met Thr Asp Val Asp Arg Asn Gly Thr  
340 345 350

355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500

Ile Gln Ala Thr Pro Ser Ser Thr Thr Glu Glu Thr Ala Thr Gln Lys  
385 390 395 400

Glu Gln Trp Phe Gly Asn Arg Trp His Glu Gly Tyr Arg Gln Thr Pro  
405 410 415

Arg Glu Asp Ser His Ser Thr Thr Gly Thr Ala Ala Ala Ser Ala His  
420 425 430

Thr Ser His Pro Met Gln Gly Arg Thr Thr Pro Ser Pro Glu Asp Ser  
435 440 445

Ser Trp Thr Asp Phe Phe Asn Pro Ile Ser His Pro Met Gly Arg Gly  
450 455 460

His Gln Ala Gly Arg Arg Met Asp Met Asp Ser Ser His Ser Thr Thr  
465 470 475 480

Leu Gln Pro Thr Ala Asn Pro Asn Thr Gly Leu Val Glu Asp Leu Asp  
485 490 495

Arg Thr Gly Pro Leu Ser Met Thr Thr Gln Gln Ser Asn Ser Gln Ser  
500 505 510

Phe Ser Thr Ser His Glu Gly Leu Glu Glu Asp Lys Asp His Pro Thr  
515 520 525

Thr Ser Thr Leu Thr Ser Ser Asn Arg Asn Asp Val Thr Gly Gly Arg  
530 535 540

Arg Asp Pro Asn His Ser Glu Gly Ser Thr Thr Leu Leu Glu Gly Tyr  
545 550 555 560

Thr Ser His Tyr Pro His Thr Lys Glu Ser Arg Thr Phe Ile Pro Val  
565 570 575

Thr Ser Ala Lys Thr Gly Ser Phe Gly Val Thr Ala Val Thr Val Gly

Phe His Pro Ser Gly Gly Ser His Thr Thr His Gly Ser Glu Ser Asp  
610 615 620

Gly His Ser His Gly Ser Gln Glu Gly Gly Ala Asn Thr Thr Ser Gly  
625 630 635 640

Pro Ile Arg Thr Pro Gln Ile Pro Glu Trp Leu Ile Ile Leu Ala Ser  
645 650 655

Leu Leu Ala Leu Ala Leu Ile Leu Ala Val Cys Ile Ala Val Asn Ser  
660 665 670

Arg Arg Arg Cys Gly Gln Lys Lys Lys Leu Val Ile Asn Ser Gly Asn  
675 680 685

Gly Ala Val Glu Asp Arg Lys Pro Ser Gly Leu Asn Gly Glu Ala Ser  
690 695 700

Lys Ser Gln Glu Met Val His Leu Val Asn Lys Glu Ser Ser Glu Thr  
705 710 715 720

Pro Asp Gln Phe Met Thr Ala Asp Glu Thr Arg Asn Leu Gln Asn Val  
725 730 735

Asp Met Lys Ile Gly Val  
740

<210> 17  
<211> 930  
<212> DNA  
<213> Homo sapiens

<220>  
<221> Gene  
<222> (1)..(930)  
<223> Tetraspan TM4SF (TSPAN-3), mRNA

<220>  
<221> CDS  
<222> (1)..(762)  
<223>

LOCUS  
TSPAN-3, 930 bp, 1988, 1988, 1988, 1988

<304> 1399  
 <305> 1  
 <306> 101-104  
 <307> 1998  
 <309>  
 <313> (1)..(930)

<400> 17

atg ggc cag tgc ggc atc acc tcc tcc aag acc gtg ctg gtc ttt etc	48
Met Gly Gln Cys Gly Ile Thr Ser Ser Lys Thr Val Leu Val Phe Leu	
1 5 10 15	

aac etc atc ttc tgg ggg gca gct ggc att tta tgc tat gtg gga gcc	96
Asn Leu Ile Phe Trp Gly Ala Ala Gly Ile Leu Cys Tyr Val Gly Ala	
20 25 30	

tat gtc ttc atc act tat gat gac tat gac cac ttc ttt gaa gat gtg	144
Tyr Val Phe Ile Thr Tyr Asp Asp Tyr Asp His Phe Phe Glu Asp Val	
35 40 45	

tac acg etc atc cct gct gta gtg atc ata gct gta gga gcc ctg ctt	192
Tyr Thr Leu Ile Pro Ala Val Val Ile Ile Ala Val Gly Ala Leu Leu	
50 55 60	

ttc atc att ggg cta att ggc tgc tgt gcc aca atc cgg gaa agt cgc	240
Phe Ile Ile Gly Leu Ile Gly Cys Cys Ala Thr Ile Arg Glu Ser Arg	
65 70 75 80	

tgt gga ctt gcc acg ttt gtc atc atc ctg etc ttg gtt ttt gtc aca	288
Cys Gly Leu Ala Thr Phe Val Ile Ile Leu Leu Leu Val Phe Val Thr	
85 90 95	

gaa gtt gtt gta gtg gtt ttg gga tat gtt tac aga gca aag gtg gaa	336
Glu Val Val Val Val Val Leu Gly Tyr Val Tyr Arg Ala Lys Val Glu	
100 105 110	

aat gag gtt gat cgc agc att cag aaa gtg tat aag acc tac aat gga	384
Asn Glu Val Asp Arg Ser Ile Gln Lys Val Tyr Lys Thr Tyr Asn Gly	
115 120 125	

acc aac cct gat gct gct agc cgg gct att gat tat gta gag aga gag	432
Thr Asn Pro Asp Ala Ala Ser Arg Ala Ile Asp Tyr Val Gln Arg Gln	
130 135 140	

ctg cat tgt tgt gga att cac aac tac tca gac tgg gaa aat aca gat	480
Leu His Cys Cys Gly Ile His Asn Tyr Ser Asp Trp Glu Asn Thr Asp	
145 150 155 160	

tgg ttc aaa gaa acc aaa aac gag agt gtc cct ctt agc tgc tgc aga	528
Trp Phe Lys Glu Thr Lys Asn Gln Ser Val Pro Leu Ser Cys Cys Arg	

tat gct gag ggg tgt gag gct cta gtt gtg aag aag cta caa gaa atc	624
Tyr Ala Glu Gly Cys Glu Ala Leu Val Val Lys Lys Leu Gln Glu Ile	
195 200 205	
atg atg cat gtg atc tgg gcc gca ctg gca ttt gca gct att cag ctg	672
Met Met His Val Ile Trp Ala Ala Leu Ala Phe Ala Ala Ile Gln Leu	
210 215 220	
ctg gcc atg ctg tgt gct tgc atc gtg ttg tgc aga agg agt aga gat	720
Leu Gly Met Leu Cys Ala Cys Ile Val Leu Cys Arg Arg Ser Arg Asp	
225 230 235 240	
cct gct tac gag ctc ctc atc act gcc gga acc tat gca tag	762
Pro Ala Tyr Glu Leu Leu Ile Thr Gly Gly Thr Tyr Ala	
245 250	
ttgacaactc ttgcctgagc tttttggtct tgttctgatt tggaaggtga attgagcagg	822
tctgctgctg ttggcctctg gagttcattt agttaagca catgtacact ggtgttggac	882
agagcagctt ggcttttcat gtgcccaact acttactact actgcgat	930

<210> 18  
 <211> 253  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> DOMAIN  
 <222> (1)..(11)  
 <223> Cytoplasmic (Potential)

<220>  
 <221> Gene  
 <222> (1)..(253)  
 <223> Tetraspanin-3

<220>  
 <221> TRANSMEM  
 <222> (12)..(32)  
 <223> Potential

<220>  
 <221> DOMAIN  
 <222> (33)..(50)  
 <223> Extracellular (Potential)

<220>  
 <221> TRANSMEM  
 <222> (51)..(71)

...  
 ... Cytoplasmic (Potential)

<220>  
 <221> TRANSMEM  
 <222> (86)..(106)  
 <223> Potential  
  
 <220>  
 <221> DOMAIN  
 <222> (107)..(212)  
 <223> Extracellular (potential)  
  
 <220>  
 <221> TRANSMEM  
 <222> (213)..(233)  
 <223> Potential  
  
 <220>  
 <221> DOMAIN  
 <222> (234)..(253)  
 <223> Cytoplasmic (potential)  
  
 <220>  
 <221> CARBOHYD  
 <222> (127)..(127)  
 <223> N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> CARBOHYD  
 <222> (152)..(152)  
 <223> N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> CARBOHYD  
 <222> (167)..(167)  
 <223> N-linked (GLCNAC...) (potential)  
  
 <220>  
 <221> CARBOHYD  
 <222> (183)..(183)  
 <223> N-linked (GLCNAC...) (potential)

<400> 13

Met Gly Gln Cys Gly Ile Thr Ser Ser Lys Thr Val Leu Val Phe Leu  
 1 5 10 15

Asn Leu Ile Phe Trp Gly Ala Ala Gly Ile Leu Cys Tyr Val Gly Ala  
 20 25 30

Tyr Thr Asn Ile Phe Ala Val Val Ile Ile Ala Val Gly Ala Leu Leu



50

55

60

Phe Ile Ile Gly Leu Ile Gly Cys Cys Ala Thr Ile Arg Glu Ser Arg  
65 70 75 80

Cys Gly Leu Ala Thr Phe Val Ile Ile Leu Leu Leu Val Phe Val Thr  
85 90 95

Glu Val Val Val Val Val Leu Gly Tyr Val Tyr Arg Ala Lys Val Glu  
100 105 110

Asn Glu Val Asp Arg Ser Ile Gln Lys Val Tyr Lys Thr Tyr Asn Gly  
115 120 125

Thr Asn Pro Asp Ala Ala Ser Arg Ala Ile Asp Tyr Val Gln Arg Gln  
130 135 140

Leu His Cys Cys Gly Ile His Asn Tyr Ser Asp Trp Glu Asn Thr Asp  
145 150 155 160

Trp Phe Lys Glu Thr Lys Asn Gln Ser Val Pro Leu Ser Cys Cys Arg  
165 170 175

Glu Thr Ala Ser Asn Cys Asn Gly Ser Leu Ala His Pro Ser Asp Leu  
180 185 190

Tyr Ala Glu Gly Cys Glu Ala Leu Val Val Lys Lys Leu Gln Glu Ile  
195 200 205

Met Met His Val Ile Trp Ala Ala Leu Ala Phe Ala Ala Ile Gln Leu  
210 215 220

Leu Gly Met Leu Cys Ala Cys Ile Val Leu Cys Arg Arg Ser Arg Asp  
225 230 235 240

Pro Ala Tyr Glu Leu Leu Ile Thr Gly Gly Thr Tyr Ala  
245 250



Ile Tyr Thr Leu Gly Tyr Ser Val Ser Leu Met Ser Leu Ala Thr Gly	
130 135 140	
agc ata att ctg tgc ctc ttc agg aag ctg cac tgc acc agg aat tac	480
Ser Ile Ile Leu Cys Leu Phe Arg Lys Leu His Cys Thr Arg Asn Tyr	
145 150 155 160	
atc cac ctg aac ctg ttc ctg tcc ttc atc ctg aga gcc atc tca gtg	528
Ile His Leu Asn Leu Phe Leu Ser Phe Ile Leu Arg Ala Ile Ser Val	
165 170 175	
ctg gtc aag gac gac gtt ctc tac tcc agc tct ggc acg ttg cac tgc	576
Leu Val Lys Asp Asp Val Leu Tyr Ser Ser Ser Gly Thr Leu His Cys	
180 185 190	
cct gac cag cca tcc tcc tgg gtg ggc tgc aag ctg agc ctg gtc ttc	624
Pro Asp Gln Pro Ser Ser Trp Val Gly Cys Lys Leu Ser Leu Val Phe	
195 200 205	
ctg cag tac tgc atc atg gcc aac ttc ttc tgg ctg ctg gtg gag ggg	672
Leu Gln Tyr Cys Ile Met Ala Asn Phe Phe Trp Leu Leu Val Glu Gly	
210 215 220	
ctc tac ctc cac acc ctc ctg gtg gcc atg ctc ccc cct aga agg tgc	720
Leu Tyr Leu His Thr Leu Leu Val Ala Met Leu Pro Pro Arg Arg Cys	
225 230 235 240	
ttc ctg gcc tac ctc ctg atc gga tgg ggc ctc ccc acc gtc tgc atc	768
Phe Leu Ala Tyr Leu Leu Ile Gly Trp Gly Leu Pro Thr Val Cys Ile	
245 250 255	
ggg gca tgg act gcg gcc agg ctc tac tta gaa gac acc ggt tgc tgg	816
Gly Ala Trp Thr Ala Ala Arg Leu Tyr Leu Glu Asp Thr Gly Cys Trp	
260 265 270	
gat aca aac gac cac agt gtg ccc tgg tgg gtc ata cga ata ccg att	864
Asp Thr Asn Asp His Ser Val Pro Trp Trp Val Ile Arg Ile Pro Ile	
275 280 285	
tta att tcc atc atc gtc aat ttt gtc ctt ttc att agt att ata cga	912
Leu Ile Ser Ile Ile Val Asn Phe Val Leu Phe Ile Ser Ile Ile Arg	
290 295 300	
att ttg ctg cag aag tta asa tcc cca gat gtc ggc ggc aac gac cag	960
Ile Leu Leu Gln Lys Leu Thr Ser Pro Asp Val Gly Gly Asn Asp Gln	
305 310 315 320	
tct cag tac aag agg ctg gcc aag tcc acg ctc ctg ctt atc ccg ctg	1008
Ser Gln Tyr Lys Arg Leu Ala Lys Ser Thr Leu Leu Leu Ile Pro Leu	
325 330 335	
gca gca gca gca gca gca gca gca gca gca gca gca gca gca gca gca	
Leu Lys Tyr Gln Ile Leu Ile Phe Leu Tyr Leu Gly Ser Phe Gln Gly	

355	360	365	
ctg gtg gtg gcc gtc ctc tac tgt ttc ctg aac agt gag gtg cag tgc			1152
Leu Val Val Ala Val Leu Tyr Cys Phe Leu Asn Ser Glu Val Gln Cys			
370	375	380	
gag ctg aag cga aaa tgg cga agc cgg tgc cgg acc ccg tcc gcg agc			1200
Glu Leu Lys Arg Lys Trp Arg Ser Arg Cys Pro Thr Pro Ser Ala Ser			
385	390	395	400
cgg gat tac agg gtc tgc ggt tcc tcc ttc tcc cac aac ggc tcg gag			1248
Arg Asp Tyr Arg Val Cys Gly Ser Ser Phe Ser His Asn Gly Ser Glu			
405	410	415	
ggc gcc ctg cag ttc cac cgc gcg tcc cga gcc cag tcc ttc ctg caa			1296
Gly Ala Leu Gln Phe His Arg Ala Ser Arg Ala Gln Ser Phe Leu Gln			
420	425	430	
acg gag acc tcg gtc atc tag			1317
Thr Glu Thr Ser Val Ile			
435			

<210> 20  
 <211> 438  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SIGNAL  
 <222> (1)..(23)  
 <223> Potential

<220>  
 <221> GENE  
 <222> (1)..(438)  
 <223> Vasoactive Intestinal Peptide Receptor 2

<220>  
 <221> CHAIN  
 <222> (24)..(438)  
 <223> Vasoactive Intestinal Polypeptide Receptor 2

<220>  
 <221> DOMAIN  
 <222> (24)..(126)  
 <223> 1 (Potential)

<220>  
 <221> TRANSMEM  
 <222> (127)..(151)

<220>  
 <221> Cytoplasmic potential

<220>  
<221> TRANSMEM  
<222> (159)..(178)  
<223> 2 (Potential)

<220>  
<221> DOMAIN  
<222> (179)..(203)  
<223> Extracellular (Potential)

<220>  
<221> TRANSMEM  
<222> (204)..(227)  
<223> 3 (Potential)

<220>  
<221> DOMAIN  
<222> (228)..(240)  
<223> Cytoplasmic (Potential)

<220>  
<221> TRANSMEM  
<222> (241)..(262)  
<223> 4 (Potential)

<220>  
<221> DOMAIN  
<222> (263)..(279)  
<223> Extracellular (Potential)

<220>  
<221> TRANSMEM  
<222> (280)..(303)  
<223> 5 (Potential)

<220>  
<221> DOMAIN  
<222> (304)..(328)  
<223> Cytoplasmic (Potential)

<220>  
<221> TRANSMEM  
<222> (329)..(348)  
<223> 6 (Potential)

<220>  
<221> DOMAIN  
<222> (349)..(360)  
<223> Extracellular (Potential)

<220>

<221> DOMAIN  
<222> (381)..(438)  
<223> Cytoplasmic (potential)

<220>  
<221> CARBOHYD  
<222> (58)..(58)  
<223> N-linked (GLCNAC...) (Potential)

<220>  
<221> CARBOHYD  
<222> (88)..(88)  
<223> N-linked (GLCNAC...) (Potential)

<220>  
<221> CARBOHYD  
<222> (92)..(92)  
<223> N-linked (GLCNAC...) (Potential)

<400> 20

Met Arg Thr Leu Leu Pro Pro Ala Leu Leu Thr Cys Trp Leu Leu Ala  
1 5 10 15

Pro Val Asn Ser Ile His Pro Glu Cys Arg Phe His Leu Glu Ile Gln  
20 25 30

Glu Glu Glu Thr Lys Cys Ala Glu Leu Leu Arg Ser Gln Thr Glu Lys  
35 40 45

His Lys Ala Cys Ser Gly Val Trp Asp Asn Ile Thr Cys Trp Arg Pro  
50 55 60

Ala Asn Val Gly Glu Thr Val Thr Val Pro Cys Pro Lys Val Phe Ser  
65 70 75 80

Asn Phe Tyr Ser Lys Ala Gly Asn Ile Ser Lys Asn Cys Thr Ser Asp  
85 90 95

Gly Trp Ser Glu Thr Phe Pro Asp Phe Val Asp Ala Cys Gly Tyr Ser  
100 105 110

Asp Pro Glu Asp Glu Ser Lys Ile Thr Phe Tyr Ile Leu Val Lys Ala

Ser Ile Ile Leu Cys Leu Phe Arg Lys Leu His Cys Thr Arg Asn Tyr  
145 150 155 160

Ile His Leu Asn Leu Phe Leu Ser Phe Ile Leu Arg Ala Ile Ser Val  
165 170 175

Leu Val Lys Asp Asp Val Leu Tyr Ser Ser Ser Gly Thr Leu His Cys  
180 185 190

Pro Asp Gln Pro Ser Ser Trp Val Gly Cys Lys Leu Ser Leu Val Phe  
195 200 205

Leu Gln Tyr Cys Ile Met Ala Asn Phe Phe Trp Leu Leu Val Glu Gly  
210 215 220

Leu Tyr Leu His Thr Leu Leu Val Ala Met Leu Pro Pro Arg Arg Cys  
225 230 235 240

Phe Leu Ala Tyr Leu Leu Ile Gly Trp Gly Leu Pro Thr Val Cys Ile  
245 250 255

Gly Ala Trp Thr Ala Ala Arg Leu Tyr Leu Glu Asp Thr Gly Cys Trp  
260 265 270

Asp Thr Asn Asp His Ser Val Pro Trp Trp Val Ile Arg Ile Pro Ile  
275 280 285

Leu Ile Ser Ile Ile Val Asn Phe Val Leu Phe Ile Ser Ile Ile Arg  
290 295 300

Ile Leu Leu Gln Lys Leu Thr Ser Pro Asp Val Gly Gly Asn Asp Gln  
305 310 315 320

Ser Gln Tyr Lys Arg Leu Ala Lys Ser Thr Leu Leu Leu Ile Pro Leu  
325 330 335

Phe Gly Val His Tyr Met Val Phe Ala Val Phe Pro Ile Ser Ile Ser  
340 345 350





tcagcggtag caacctcgcc ccttgcaaca aaggcagact gagcgccaga gaggacgttt	240
ccaactcaaa a atg cag gct caa cag tac cag cag cag cgt cga aaa ttt	290
Met Gln Ala Gln Gln Tyr Gln Gln Gln Arg Arg Lys Phe	
1 5 10	
gca gct gcc ttc ttg gca ttc att ttc ata ctg gca gct gtg gat act	338
Ala Ala Ala Phe Leu Ala Phe Ile Phe Ile Leu Ala Ala Val Asp Thr	
15 20 25	
gct gaa gca ggg aag aaa gag aaa cca gaa aaa aaa gtg aag aag tct	386
Ala Glu Ala Gly Lys Lys Glu Lys Pro Glu Lys Lys Val Lys Lys Ser	
30 35 40 45	
gac tgt gga gaa tgg cag tgg agt gtg tgt gtg ccc acc agt gga gac	434
Asp Cys Gly Glu Trp Gln Trp Ser Val Cys Val Pro Thr Ser Gly Asp	
50 55 60	
tgt ggg ctg ggc aca cgg gag ggc act cgg act gga gct gag tgc aag	482
Cys Gly Leu Gly Thr Arg Glu Gly Thr Arg Thr Gly Ala Glu Cys Lys	
65 70 75	
caa acc atg aag acc cag aga tgt aag atc ccc tgc aac tgg aag aag	530
Gln Thr Met Lys Thr Gln Arg Cys Lys Ile Pro Cys Asn Trp Lys Lys	
80 85 90	
caa ttt ggc gcg gag tgc aaa tac cag ttc cag gcc tgg gga gaa tgt	578
Gln Phe Gly Ala Glu Cys Lys Tyr Gln Phe Gln Ala Trp Gly Glu Cys	
95 100 105	
gac ctg aac aca gcc ctg aag acc aga act gga agt ctg aag cga gcc	626
Asp Leu Asn Thr Ala Leu Lys Thr Arg Thr Gly Ser Leu Lys Arg Ala	
110 115 120 125	
ctg cac aat gcc gaa tgc cag aag act gtc acc atc tcc aag ccc tgt	674
Leu His Asn Ala Glu Cys Gln Lys Thr Val Thr Ile Ser Lys Pro Cys	
130 135 140	
ggc aaa ctg acc aag ccc aaa cct caa gca gaa tct aag aag aag aaa	722
Gly Lys Leu Thr Lys Pro Lys Pro Gln Ala Glu Ser Lys Lys Lys Lys	
145 150 155	
aag gaa ggc aag aaa cag gag aag atg ctg gat taa aagatgtcac	768
Lys Glu Gly Lys Lys Gln Glu Lys Met Leu Asp	
160 165	
ctgtggaaca taaaaaggac atcagcaaac aggatcagtt aactattgca tttatatgta	828
cctgtaggett tgtatttcaaa aattatctat agctaagtac acaataagca aaaacaaaaa	888
g	889

<220>  
 <221> SIGNAL  
 <222> (1)..(32)  
 <223>

<220>  
 <221> GENE  
 <222> (1)..(168)  
 <223> Pleiotrophin

<220>  
 <221> CHAIN  
 <222> (33)..(168)  
 <223> Pleiotrophin

<220>  
 <221> DISULFID  
 <222> (47)..(76)  
 <223> By similarity.

<220>  
 <221> DISULFID  
 <222> (55)..(85)  
 <223> By similarity.

<220>  
 <221> DISULFID  
 <222> (62)..(89)  
 <223> By similarity.

<220>  
 <221> DISULFID  
 <222> (99)..(131)  
 <223> By similarity.

<220>  
 <221> DISULFID  
 <222> (109)..(141)  
 <223> By similarity.

<400> 22

Met Gln Ala Gln Gln Tyr Gln Gln Gln Arg Arg Lys Phe Ala Ala Ala  
 1 5 10 15

Phe Leu Ala Phe Ile Phe Ile Leu Ala Ala Val Asp Thr Ala Glu Ala  
 20 25 30

Gln Arg Gln Arg Gln Val Tyr Val Ser Thr Ser Gly Asp Tyr Gly Leu

50

55

60

Gly Thr Arg Glu Gly Thr Arg Thr Gly Ala Glu Cys Lys Gln Thr Met  
65 70 75 80

Lys Thr Gln Arg Cys Lys Ile Pro Cys Asn Trp Lys Lys Gln Phe Gly  
85 90 95

Ala Glu Cys Lys Tyr Gln Phe Gln Ala Trp Gly Glu Cys Asp Leu Asn  
100 105 110

Thr Ala Leu Lys Thr Arg Thr Gly Ser Leu Lys Arg Ala Leu His Asn  
115 120 125

Ala Glu Cys Gln Lys Thr Val Thr Ile Ser Lys Pro Cys Gly Lys Leu  
130 135 140

Thr Lys Pro Lys Pro Gln Ala Glu Ser Lys Lys Lys Lys Lys Glu Gly  
145 150 155 160

Lys Lys Gln Glu Lys Met Leu Asp  
165

<210> 23  
<211> 3143  
<212> DNA  
<213> Homo sapiens

<220>  
<221> Gene  
<222> (1)..(3143)  
<223> Osteopontin

<220>  
<221> CDS  
<222> (374)..(427)  
<223>

<220>  
<221> CDS  
<222> (537)..(575)  
<223>

<221> CDS  
<222> (1232)..(1273)  
<223>

<220>  
<221> CDS  
<222> (1597)..(1920)  
<223>

<220>  
<221> CDS  
<222> (2305)..(2709)  
<223>

<300>  
<301> Crosby, A.H., Edwards, S.J., Murray, J.C. and Dixon, M.J.  
<302> Genomic organization of the human osteopontin gene: exclusion of the locus from a causative role in the pathogenesis of dentinogenesis imperfecta type II  
<303> Genomics  
<304> 27  
<305> 1  
<306> 155-160  
<307> 1995  
<309>  
<313> (1)..(3143)

<400> 23  
ggggaagtgt gggagcaggt gggctgggca gtggcagaaa cctgatgaca caatctcgcc 60  
gcctccctgt gttggtggag gatgtctgca gcagcattta aattctggga gggcttggtt 120  
gtcagcagca gcaggaggag gcagagacag catcgtcggg accagactcg tctcaggcca 180  
gttgcaacct tctcagccaa acgcccacca aggtacagct tcagtttgct actgggttgt 240  
gcattcagct gaatttcatt gggaagtcca aattctaagg aaaaaaatgt ggtagtataa 300  
aaaggtatca ctgttgtaac ctatgaagat gtcagctatt cctttgaaat attttgcagg 360  
aaaaactcact acc atg aga att gca gtg att tgc ttt tgc ctc cta ggc 409  
Met Arg Ile Ala Val Ile Cys Phe Cys Leu Leu Gly  
1 5 10  
atc acc tgt gcc ata cca gtgagtagag ttgcatttta aagaaaattc 457  
Ile Thr Cys Ala Ile Pro  
15  
ctgaaaataa ctgaattata tctttccata tctcaggaga acattcttct aattcttctt 517

... .. 528

Lys Gln  
30

cgagaggtgc aagaaacgta tttgctgcga tcaaatagagt tcatatttgt aaagcaattt 685

gaaagagtgc ctagccacaca gtaagtgcga cataagagtt tgttaaatga atctgcaaaa 745

aaaaaaaaaa ttacaaaaag gtacctaagg gtccgggtga ctatatgctt ccatcaagac 805

tagtgaagaa tgggtgtttt ttccattcat ccctacattt ctttttttaa taatgataaa 865

catgcaactt tttttagt ctt tac aac aaa tac cca gat gct gtg gcc aca 916  
Leu Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr  
35 40

tgg cta aac cct gac cca tct cag aag cag aat ctc cta gcc cca cag 964  
Trp Leu Asn Pro Asp Pro Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln  
45 50 55

gtatttttaa acttttcata attaaactac agtgatgaaa gatagccaca ctcaggccat 1024

ttgggtgtct cagatgaatc ctgccctgcc tgctggcaaa catgtgctta ggacattgac 1084

tgatctgcca tgttggtctt tctctgtgtt aagccatcca cagatgaggc tgaaaaataa 1144

aaactgcttt ggattaaaaa ggttaacttt tgaataaaaa agctaggcat gtgtgatgcy 1204

cactaacag tgccattcct tcttcag aat gct gtg tcc tct gaa gaa acc aat 1258  
Asn Ala Val Ser Ser Glu Glu Thr Asn  
60 65

gac ttt aaa caa gag gtaagttctc attttcaatc agaggcccat catgccttga 1313  
Asp Phe Lys Gln Glu  
70

agagatgaaa gaaggcattg cctggattct cttctgatga aatttcatta gcaagttttc 1373

cagctaattg gcagtctaaa acttgctcat aaataaaaca tgtatttact aaatatcaga 1433

aatactaggt ttctcggat aacctaaaag ccatgggatg tactgtgaat gcaaagattc 1493

tgaaactaaa taaaaagaaa gatagtaaaa gactaatgtg ctataaaggc taagggaaaa 1553

taaaaaccca tatattaatt ttccggcca tcttaatttt cag acc ctt cca agt 1608  
Thr Leu Pro Ser  
75

aag tcc aac gaa agc cat gac cac atg gat gat atg gat gat gaa gat 1656  
Lys Ser Asn Glu Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp  
80 85 90

Arg Arg Thr Arg Arg Thr Arg Arg Pro His Ala Ser Arg Glu Ser His

110	115	120	
cat tct gat gaa tct gat gaa ctg gtc act gat ttt ccc acg gac ctg			1800
His Ser Asp Glu Ser Asp Glu Leu Val Thr Asp Phe Pro Thr Asp Leu			
125	130	135	140
cca gca acc gaa gtt ttc act cca gtt gtc ccc aca gta gac aca tat			1848
Pro Ala Thr Glu Val Phe Thr Pro Val Val Pro Thr Val Asp Thr Tyr			
	145	150	155
gat ggc cga ggt gat agt gtg gtt tat gga ctg agg tca aaa tct aag			1896
Asp Gly Arg Gly Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys			
	160	165	170
aag ttt cgc aga cct gac atc cag gtaaactcct taacagacac acctgatggt			1950
Lys Phe Arg Arg Pro Asp Ile Gln			
	175	180	
tctgactago gctcaagtct aggaaaccac agtttgcata ttcattcatt cattcatcca			2010
ttcattcato cattcagcaa gaattcattc atattctact ttatgaccat tgaatacaaaa			2070
tctttttctg cttggcggtt ttgtgaagtc tacataatct ctctctagat ttgattctca			2130
aacacaatct tactttttga aatcctggat caaagtaaca tgctagtatt atttcagcca			2190
gatttagaca atttttagta taagatgacc taaaagctag agagtggaaa aggattacca			2250
tattcccatc cctagccgtt catataatta ttcttcattt gtgcgctgat tcag tac			2307
			Tyr
cct gat gct aca gac gag gac atc acc tca cac atg gaa agc gag gag			2355
Pro Asp Ala Thr Asp Glu Asp Ile Thr Ser His Met Glu Ser Glu Glu			
	185	190	195
ttg aat ggt gca tac aag gcc atc ccc gtt gcc cag gac ctg aac gcg			2403
Leu Asn Gly Ala Tyr Lys Ala Ile Pro Val Ala Gln Asp Leu Asn Ala			
	200	205	210
cct tct gat tgg gac agc cgt ggg aag gac agt tat gaa acg agt cag			2451
Pro Ser Asp Trp Asp Ser Arg Gly Lys Asp Ser Tyr Gln Thr Ser Gln			
	215	220	225
ctg gat gac cag agt gct gaa acc cac agc cac aag cag tcc aga tta			2499
Leu Asp Asp Gln Ser Ala Glu Thr His Ser His Lys Gln Ser Arg Leu			
	235	240	245
tat aag cgg aaa gcc aat gat gag agc aat gag cat tcc gat gtg att			2547
Tyr Lys Arg Lys Ala Asn Asp Glu Ser Asn Glu His Ser Asp Val Ile			
	250	255	260

Phe His Ser His Glu Asp Met Leu Val Val Asp Pro Lys Ser Lys Glu  
 280 285 290

gaa gat aaa cac ctg aaa ttt cgt att tct cat gaa tta gat agt gca 2691  
 Glu Asp Lys His Leu Lys Phe Arg Ile Ser His Glu Leu Asp Ser Ala  
 295 300 305

tct tct gag gtc aat taa aaggagaaaa aatacaatctt ctcactttgc 2739  
 Ser Ser Glu Val Asn  
 310

atttagtcaa aagaaaaaat gctttatagc aaaatgaaag agaacatgaa atgcttcttt 2799

ctcagtttat tgggttgatg tgtatctatt tgagtctgga aataactaat gtgtttgata 2859

attagtttag tttgtggctt catggaaact ccctgtaaac aaaagcttca gggttatgtc 2919

tatgttcatt ctatagaaga aatgcaaact atcactgtat tttaatatctt gttattctct 2979

catgaataga aatttatgta gaagcaaaca aaatactttt acccacttaa aaagagaata 3039

taacatttta tgtcactata atcttttgtt ttttaagtta gtgtatatatt tgttgtgatt 3099

atcttttgtg gtgtgaataa atcttttatc ttgaatgtaa taag 3143

<210> 24  
 <211> 314  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SIGNAL  
 <222> (1)..(16)  
 <223> Potential

<220>  
 <221> Gene  
 <222> (1)..(314)  
 <223> Osteopontin

<220>  
 <221> CHAIN  
 <222> (17)..(314)  
 <223> Osteopontin

<220>  
 <221> SITE  
 <222> (159)..(161)  
 <223> Cell attachment site

..

<221> CARBOHYD  
<222> (106)..(106)  
<223> N-linked (GLCNAC...) (Potential)

<220>  
<221> VARSPLIC  
<222> (31)..(57)  
<223> Missing (In Isoform C)

<220>  
<221> VARSPLIC  
<222> (58)..(71)  
<223> Missing (In Isoform B)

<400> 24

Met Arg Ile Ala Val Ile Cys Phe Cys Leu Leu Gly Ile Thr Cys Ala  
1 5 10 15

Ile Pro Val Lys Gln Ala Asp Ser Gly Ser Ser Glu Glu Lys Gln Leu  
20 25 30

Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr Trp Leu Asn Pro Asp Pro  
35 40 45

Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln Asn Ala Val Ser Ser Glu  
50 55 60

Glu Thr Asn Asp Phe Lys Gln Glu Thr Leu Pro Ser Lys Ser Asn Glu  
65 70 75 80

Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp Asp Asp Asp His  
85 90 95

Val Asp Ser Gln Asp Ser Ile Asp Ser Asn Asp Ser Asp Asp Val Asp  
100 105 110

Asp Thr Asp Asp Ser His Gln Ser Asp Glu Ser His His Ser Asp Glu  
115 120 125

Ser Asp Glu Leu Val Thr Asp Phe Pro Thr Asp Leu Pro Ala Thr Glu  
130 135 140



Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys Lys Phe Arg Arg  
165 170 175

Pro Asp Ile Gln Tyr Pro Asp Ala Thr Asp Glu Asp Ile Thr Ser His  
180 185 190

Met Glu Ser Glu Glu Leu Asn Gly Ala Tyr Lys Ala Ile Pro Val Ala  
195 200 205

Gln Asp Leu Asn Ala Pro Ser Asp Trp Asp Ser Arg Gly Lys Asp Ser  
210 215 220

Tyr Glu Thr Ser Gln Leu Asp Asp Gln Ser Ala Glu Thr His Ser His  
225 230 235 240

Lys Gln Ser Arg Leu Tyr Lys Arg Lys Ala Asn Asp Glu Ser Asn Glu  
245 250 255

His Ser Asp Val Ile Asp Ser Gln Glu Leu Ser Lys Val Ser Arg Glu  
260 265 270

Phe His Ser His Glu Phe His Ser His Glu Asp Met Leu Val Val Asp  
275 280 285

Pro Lys Ser Lys Glu Glu Asp Lys His Leu Lys Phe Arg Ile Ser His  
290 295 300

Glu Leu Asp Ser Ala Ser Ser Glu Val Asn  
305 310

<210> 25

<211> 259

<212> PET

<213> Homo sapiens

<220>

<221> Gene

<222> (1)..(259)

<223> Carbonic Anhydrase domain of human carbonic anhydrase III

Glu Leu Phe Pro Asn Ala Lys Gly Glu Asn Gln Ser Pro Ile Glu Leu  
20 25 30

His Thr Lys Asp Ile Arg His Asp Pro Ser Leu Gln Pro Trp Ser Val  
35 40 45

Ser Tyr Asp Gly Gly Ser Ala Lys Thr Ile Leu Asn Asn Gly Lys Thr  
50 55 60

Cys Arg Val Val Phe Asp Asp Thr Tyr Asp Arg Ser Met Leu Arg Gly  
65 70 75 80

Gly Pro Leu Pro Gly Pro Tyr Arg Leu Arg Gln Phe His Leu His Trp  
85 90 95

Gly Ser Ser Asp Asp His Gly Ser Glu His Thr Val Asp Gly Val Lys  
100 105 110

Tyr Ala Ala Glu Leu His Leu Val His Trp Asn Pro Lys Tyr Asn Thr  
115 120 125

Phe Lys Glu Ala Leu Lys Gln Arg Asp Gly Ile Ala Val Ile Gly Ile  
130 135 140

Phe Leu Lys Ile Gly His Glu Asn Gly Glu Phe Gln Ile Phe Leu Asp  
145 150 155 160

Ala Leu Asp Lys Ile Lys Thr Lys Gly Lys Glu Ala Pro Phe Thr Lys  
165 170 175

Phe Asp Pro Ser Cys Leu Phe Pro Ala Cys Arg Asp Tyr Trp Thr Tyr  
180 185 190

Gln Gly Ser Phe Thr Thr Pro Pro Cys Glu Glu Cys Ile Val Trp Leu  
195 200 205

Leu Leu Lys Glu Pro Met Thr Val Ser Ser Asp Gln Met Ala Lys Leu  
210 215 220

Ser Asn Thr Asn Ile Thr Gln Pro Ile Asn Asn Arg Val Val Asn Ala

245

250

255

Ser Phe Lys

&lt;210&gt; 26

&lt;211&gt; 260

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; Gene

&lt;222&gt; (1)..(260)

&lt;223&gt; Carbonic anhydrase domain of human carbonic anhydrase I

&lt;400&gt; 26

Ala	Ser	Pro	Asp	Trp	Gly	Tyr	Asp	Asp	Lys	Asn	Gly	Pro	Glu	Gln	Trp
1				5					10					15	

Ser	Lys	Leu	Tyr	Pro	Ile	Ala	Asn	Gly	Asn	Asn	Gln	Ser	Pro	Val	Asp
		20						25					30		

Ile	Lys	Thr	Ser	Glu	Thr	Lys	His	Asp	Thr	Ser	Leu	Lys	Pro	Ile	Ser
		35					40					45			

Val	Ser	Tyr	Asn	Pro	Ala	Thr	Ala	Lys	Glu	Ile	Ile	Asn	Val	Gly	His
	50					55					60				

Ser	Phe	His	Val	Asn	Phe	Glu	Asp	Asn	Asp	Asn	Arg	Ser	Val	Leu	Lys
65					70					75				80	

Gly	Gly	Pro	Phe	Ser	Asp	Ser	Tyr	Arg	Leu	Phe	Gln	Phe	His	Phe	His
			35						90					95	

Trp	Gly	Ser	Thr	Asn	Glu	His	Gly	Ser	Glu	His	Thr	Val	Asp	Gly	Val
			100					105						110	

Lys	Tyr	Ser	Ala	Glu	Leu	His	Val	Ala	His	Trp	Asn	Ser	Ala	Lys	Tyr
		115					120					125			

Gly	Val	Leu	Met	Lys	Val	Gly	His	Ala	Asn	Ile	Lys	Leu	Gln	Lys	Val
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

145 150 155 160

Leu Asp Ala Leu Gln Ala Ile Lys Thr Lys Gly Lys Arg Ala Pro Phe  
165 170 175

Thr Asn Phe Asp Pro Ser Thr Leu Leu Pro Ser Ser Leu Asp Phe Trp  
180 185 190

Thr Tyr Pro Gly Ser Leu Thr His Pro Pro Leu Tyr Glu Ser Val Thr  
195 200 205

Trp Ile Ile Cys Lys Glu Ser Ile Ser Val Ser Ser Glu Gln Leu Ala  
210 215 220

Gln Phe Arg Ser Leu Leu Ser Asn Val Glu Gly Asp Asn Ala Val Pro  
225 230 235 240

Met Gln His Asn Asn Arg Pro Thr Gln Pro Leu Lys Gly Arg Thr Val  
245 250 255

Arg Ala Ser Phe  
260

<210> 27  
<211> 337  
<212> PRT  
<213> Homo sapiens

<220>  
<221> Gene  
<222> (1)..(337)  
<223> Carbonic anhydrase domain of human carbonic anhydrase VIX

400 27

Met Leu Phe Ser Ala Leu Leu Leu Glu Val Ile Trp Ile Leu Ala Ala  
1 5 10 15

Asp Gly Gly Gln His Trp Thr Tyr Glu Gly Pro His Gly Gln Asp His  
20 25 30

Asp Ile Gln Thr Arg Ser Val Thr His Arg Ile Arg Leu Ile Ala Leu

50

55

60

Gln Pro His Gly Tyr Asp Gln Pro Gly Thr Glu Pro Leu Asp Leu His  
65 70 75 80

Asn Asn Gly His Thr Val Gln Leu Ser Leu Pro Ser Thr Leu Tyr Leu  
85 90 95

Gly Gly Leu Pro Arg Lys Tyr Val Ala Ala Gln Leu His Leu His Trp  
100 105 110

Gly Gln Lys Gly Ser Pro Gly Gly Ser Glu His Gln Ile Asn Ser Glu  
115 120 125

Ala Thr Phe Ala Glu Leu His Ile Val His Tyr Asp Ser Asp Ser Tyr  
130 135 140

Asp Ser Leu Ser Glu Ala Ala Glu Arg Pro Gln Gly Leu Ala Val Leu  
145 150 155 160

Gly Ile Leu Ile Glu Val Gly Glu Thr Lys Asn Ile Ala Tyr Glu His  
165 170 175

Ile Leu Ser His Leu His Glu Val Arg His Lys Asp Gln Lys Thr Ser  
180 185 190

Val Pro Pro Phe Asn Leu Arg Glu Leu Leu Pro Lys Gln Leu Gly Gln  
195 200 205

Tyr Phe Arg Tyr Asn Gly Ser Leu Thr Thr Pro Pro Cys Tyr Gln Ser  
210 215 220

Val Leu Trp Thr Val Phe Tyr Arg Arg Ser Gln Ile Ser Met Glu Gln  
225 230 235 240

Leu Glu Lys Leu Gln Gly Thr Leu Phe Ser Thr Glu Glu Glu Pro Ser  
245 250 255

Met Thr Phe Ala Ser His Thr Gln Asn Gly Ser Ser Tyr Leu Thr Ser  
260 265 270

Glu Met Leu Ser Leu Gly Val Gly Ile Leu Val Gly Cys Leu Cys Leu  
290 295 300

Leu Leu Ala Val Tyr Phe Ile Ala Arg Lys Ile Arg Lys Lys Arg Leu  
305 310 315 320

Glu Asn Arg Lys Ser Val Val Phe Thr Ser Ala Gln Ala Thr Thr Glu  
325 330 335

Ala

<210> 28  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 28  
cagcagttgg atggaagagg ac 22

<210> 29  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 29  
cactgagatt ctggcactat tc 22

<210> 30  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<210> 31  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 31  
ttgactggct caggagtata g 21

<210> 32  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 32  
ctgataatga gggctcccaa c 21

<210> 33  
<211> 24  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 33  
ctctgcactt cctggtaaaa ctct 24

<210> 34  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer

<400> 34  
cagcagttgg atggaagagg ac 22

Artificial primer

<220>

<223> Primer

<400> 35

ctctgcactt cctggtaaaa ctct

24